Б. П. ДЕМИДОВИЧ, И. А. МАРОН, Э. З. ШУВАЛОВА

ЧИСЛЕННЫЕ МЕТОДЫ АНАЛИЗА

ПРИБЛИЖЕНИЕ ФУНКЦИЙ, ДИФФЕРЕНЦИАЛЬНЫЕ И ИНТЕГРАЛЬНЫЕ УРАВНЕНИЯ

Под редакцией Б. П. ДЕМИДОВИЧА

ИЗДАНИЕ ТРЕТЬЕ, ПЕРЕРАБОТАННОЕ

Допущено Министерством высшего и среднего специального образования СССР в качестве учебного пособия для студентов высших технических учебных заведений

ИЗДАТЕЛЬСТВО «НАУКА»
ГЛАВНАЯ РЕДАКЦИЯ
ФИЗИКО-МАТЕМАТИЧЕСКОЙ ЛИТЕРАТУРЫ
МОСКВА 1967

518
Д 30
УДК 518

АННОТАЦИЯ

В книге излагаются избранные вопросы вычислительной математики, и по содержанию она является продолжением учебного пособия Б. П. Демидовича и И. А. Марона «Основы вычислительной математики».

Настоящее, третье издание отличается от предыдущего более доходчивым изложением. Добавлены новые примеры.

Рассчитана на студентов технических, экономических и педагогических институтов. Может быть использована также инженерами, вычислителями и лицами, работающими в области прикладной математики.

Приближение функций, дифференциальные и интегральные уравнения

М., 1967 г., 368 стр. с илл.

Редактор *М. М. Горячая*

Техн. редактор *К. Ф. Брудно* Корректор *О. А. Бутусо*

Сдано в набор 7/II 1967 г. Подписано к печати 16/VI 1967 г. Бумага 60×90/1
Физ. печ. л. 23. Условн. печ. л. 23. Уч.-изд. л. 24,09. Тираж 75 000 экз. Т-0698
Цена книги 94 коп. Заказ № 1389.

Издательство «Наука»
Главная редакция физико-математической литературы
Москва, В-71, Ленинский проспект, 15.

Ордена Трудового Красного Знамени
Первая Образцовая типография имени А. А. Жданова
Главполиграфпрома Комитета по печати при Совете Министров СССР
Москва, Ж-54, Валовая, 28.

2-2
6-6

ОГЛАВЛЕНИЕ

Из предисловия к первому изданию 6

Предисловие ко второму изданию 8

Предисловие к третьему изданию 8

Введение 9

Литература к введению 11

Глава I. **Приближение функций** 12

§ 1. Постановка задачи о приближении функций 12
§ 2. Интерполирование функций 13
§ 3. Интерполирование периодических функций с помощью тригонометрических полиномов 17
§ 4. Точечное квадратичное аппроксимирование функций 21
§ 5. Функции, ортогональные на точечном множестве 27
§ 6. Полиномы Чебышева, ортогональные на системе равноотстоящих точек 34
§ 7. Интегральное квадратичное аппроксимирование функций на отрезке 40
§ 8. Ортогональные на промежутке системы функций 43
§ 9. Понятие о гармоническом анализе 49
§ 10. Полиномы Лежандра 56
§ 11. Ортогональность с весом 63
§ 12. Полиномы Чебышева 65
§ 13. Понятие о равномерном приближении функций 71
Литература к главе I 78

Глава II. **Эмпирические формулы** 79

§ 1. Вводные замечания 79
§ 2. Линейная зависимость 82
§ 3. Метод выравнивания 84
§ 4. Квадратичная (параболическая) зависимость 89
§ 5. Определение параметров эмпирической формулы 91
§ 6. Метод выбранных точек 92
§ 7. Метод средних 93
§ 8. Метод наименьших квадратов 96
§ 9. Некоторые соображения о выборе вида эмпирической формулы с двумя параметрами 101
§ 10. Эмпирические формулы, содержащие три параметра 107

§ 11. Уточнение полученной эмпирической формулы 112
§ 12. Общий метод определения параметров эмпирической формулы 114
Литература к главе II 120

Глава III. **Приближенное решение обыкновенных дифференциальных уравнений** 121

§ 1. Общие замечания 121
§ 2. Интегрирование дифференциальных уравнений с помощью степенных рядов 128
§ 3. Метод последовательных приближений 134
§ 4. Метод численного интегрирования 140
§ 5. Метод Эйлера 144
§ 6. Модификации метода Эйлера 147
§ 7. Метод Рунге—Кутта 151
§ 8. Метод Адамса 156
§ 9. Метод А. Н. Крылова последовательных сближений . . . 163
§ 10. Метод Милна 168
§ 11. Методы, основанные на применении производных высших порядков 181
§ 12. Численное интегрирование дифференциальных уравнений второго порядка 187
§ 13. Метод Чаплыгина 191
§ 14. Метод Ньютона—Канторовича 201
§ 15. Некоторые замечания об оценке погрешностей решений дифференциальных уравнений 202
Литература к главе III 207

Глава IV. **Краевые задачи для обыкновенных дифференциальных уравнений** 209

§ 1. Общая постановка краевой задачи 209
§ 2. Линейная краевая задача 212
§ 3. Редукция к задаче Коши двухточечной краевой задачи для линейного уравнения второго порядка 217
§ 4. Метод конечных разностей 219
§ 5. Метод прогонки 224
§ 6. Метод коллокации 232
§ 7. Метод наименьших квадратов 234
§ 8. Метод Галеркина 237
§ 9. Понятие о приближенных методах решения общей краевой задачи 240
Литература к главе IV 243

Глава V. **Приближенные методы решения краевых задач для дифференциальных уравнений с частными производными** . . . 244

§ 1. Классификация дифференциальных уравнений с частными производными 244
§ 2. Начальные и краевые условия. Задача Коши. Смешанная задача. Корректность постановки смешанной задачи 247
§ 3. Краевые задачи для уравнений эллиптического типа 253
§ 4. Некоторые сведения о гармонических функциях. Единственность решения задачи Дирихле 255
§ 5. Уравнение Лапласа в конечных разностях 257
§ 6. Решение задачи Дирихле методом сеток 261
§ 7. Процесс Либмана 264
§ 8. Понятие о решении задачи Дирихле методом моделирования 270
§ 9. Понятие о решении задачи Дирихле методом Монте-Карло 272

§ 10. Метод сеток для уравнения параболического типа 278
§ 11. Устойчивость конечно-разностной схемы для решения уравнения теплопроводности 281
§ 12. Метод прогонки для уравнения теплопроводности 285
§ 13. Метод сеток для уравнений гиперболического типа 290
§ 14. Понятие о методе прямых 293
§ 15. Метод прямых для уравнения Пуассона 297
Литература к главе V 302

Глава VI. **Вариационные методы решения краевых задач** 304
§ 1. Понятие о функционале и операторе 304
§ 2. Вариационная задача 308
§ 3. Основные теоремы вариационного метода решения краевых задач 309
§ 4. Сведение линейной краевой задачи для обыкновенного дифференциального уравнения второго порядка к вариационной задаче 312
§ 5. Краевые задачи для уравнений Пуассона и Лапласа 317
§ 6. Идея метода Ритца 321
§ 7. Метод Ритца для простейшей краевой задачи 322
§ 8. Приложение метода Ритца к решению краевой задачи Штурма—Лиувилля 324
§ 9. Метод Ритца для задачи Дирихле 328
Литература к главе VI 331

Глава VII. **Интегральные уравнения** 332
§ 1. Основные виды линейных интегральных уравнений 332
§ 2. Связь между дифференциальными уравнениями и уравнениями Вольтерра 335
§ 3. Связь линейной краевой задачи с интегральным уравнением Фредгольма 337
§ 4. Метод последовательных приближений 338
§ 5. Решение интегрального уравнения методом конечных сумм 341
§ 6. Метод вырожденных ядер 345
§ 7. Метод коллокации 353
§ 8. Метод наименьших квадратов 356
§ 9. Метод моментов 358
Литература к главе VII 361

Приложение I. Ортогональные полиномы Чебышева для $n+1$ равноотстоящих точек ($\hat{P}_k(t)=\hat{P}_k(0)\,P_{k,n}(t)$) 362
Приложение II. Первые 10 полиномов Лежандра $P_n(x)$ 364
Приложение III. Первые 12 полиномов Чебышева $T_n(x)$ 364
Предметный указатель 365

ИЗ ПРЕДИСЛОВИЯ К ПЕРВОМУ ИЗДАНИЮ

В связи с потребностями новой техники инженерная практика наших дней все чаще и чаще встречается с математическими задачами, точное решение которых весьма сложно или неизвестно. В этих случаях обычно прибегают к тем или иным приближенным вычислениям. Вот почему приближенные и численные методы математического анализа получили за последние годы широкое развитие и приобрели исключительно важное значение.

Рост производительных сил в XX столетии обусловил решительный прогресс в области вычислительной техники, приведший к созданию современных электронных вычислительных машин с программным управлением. Это неограниченно расширило вычислительные возможности математики: задачи, для решения которых при ручном счете требовались годы, сейчас сплошь и рядом решаются за несколько часов, причем непосредственный счет занимает минуты.

В свою очередь, новые вычислительные средства вызвали переоценку известных методов решения задач с точки зрения целесообразности их реализации на современных вычислительных машинах и стимулировали создание более эффективных приемов.

Умелое применение вычислительной техники немыслимо без знания вычислительной математики. В настоящее время трудно себе представить творчески работающего инженера-исследователя или специалиста по экономическому планированию, не владеющего методами приближенного анализа. Массовое появление вычислительных центров, как самостоятельных, так и при ряде учебных и научно-исследовательских институтов, также неизбежно ставит вопрос о необходимости повышения математической подготовки инженеров, в первую очередь в области приближенных вычислений.

Указанные выше обстоятельства делают актуальным написание учебных пособий по вычислительной математике для инженеров, экономистов и т. д.

Настоящая книга посвящена избранным вопросам численного анализа: приближению функций и приближенному решению дифференциальных уравнений (обыкновенных и с частными производными). Такой выбор материала обусловлен тем, что вопросы, связанные с

решением алгебраических уравнений и численными методами линейной алгебры и др., имеются в вышедшей в 1960 г. книге авторов «Основы вычислительной математики» *).

Цель этой книги, как и указанной выше, дать систематическое и современное изложение важнейших приемов приближенного и численного анализа (в пределах рассматриваемых тем) на базе общего втузовского курса высшей математики.

Для расширения математического кругозора инженера дается понятие о нетрадиционных методах вычислений: методе Монте-Карло и методе моделирования.

Как и в первой книге, основные методы доведены до численных приложений: даны расчетные схемы и приведены числовые примеры с подробным ходом решения. В целях доходчивости большинство примеров рассматривается в упрощенной трактовке и носит иллюстративный характер.

Для понимания основного текста книги достаточно знания высшей математики в объеме двух первых курсов втузов машиностроительных специальностей. Необходимые сведения по математике, не входящие в общую программу втузов, излагаются в соответствующих главах. Использованная и дополнительная литература указана после каждой главы.

Книга предназначена для студентов втузов с повышенной программой по высшей математике и инженеров, занимающихся прикладными вопросами, а также для работников вычислительных бюро и центров. Кроме того, книга окажется полезной студентам физико-математических факультетов педагогических институтов и студентам экономических вузов.

В задачу авторов не входило изложение сведений по технике решения инженерных задач на электронных вычислительных машинах и программированию. По этим вопросам следует обратиться к специальным руководствам.

Авторы приносят благодарность коллективу кафедры высшей математики Артиллерийской инженерной академии им. Ф. Э. Дзержинского, принимавшему участие в обсуждении рукописи. Авторы выражают также искреннюю признательность за обстоятельные рецензии проф. Б. М. Левитану и проф. Х. Л. Смолицкому, критические замечания которых были учтены при окончательном редактировании текста.

---

*) Второе и третье издания указанной книги вышли соответственно в 1963 и 1966 гг.

Москва 1961 г. *Авторы*

ПРЕДИСЛОВИЕ КО ВТОРОМУ ИЗДАНИЮ

Во втором издании исправлены замеченные ошибки и добавлена VII глава «Интегральные уравнения». Кроме того, в конце введения и глав IV и V сделаны некоторые замечания.

Москва, 1963 г. *Авторы*

ПРЕДИСЛОВИЕ К ТРЕТЬЕМУ ИЗДАНИЮ

В третьем издании книга подверглась некоторой переработке, цель которой — сделать ее более доступной для лиц, имеющих лишь втузовское образование. Для этого изъяты или частично сокращены некоторые сложные теоретические обоснования методов и увеличено число примеров, иллюстрирующих теорию.

Наибольшие изменения внесены в главу I. Здесь написаны наново § 3 «Интерполирование периодических функций с помощью тригонометрических полиномов», § 5 «Функции, ортогональные на точечном множестве»; изменен § 6 «Полиномы Чебышева, ортогональные на системе равноотстоящих точек»; изъят параграф «Понятие о приближенном построении полинома наилучшего равномерного приближения». В конце книги дополнительно помещены приложения и предметный указатель.

Авторы просят читателей о всех своих пожеланиях и замечаниях сообщать по адресу: Москва, В-71, Ленинский проспект, 15, издательство «Наука», Главная редакция физико-математической литературы.

Москва, 1967 г. *Авторы*

ВВЕДЕНИЕ

Значение вычислительных машин и главным образом электронных вычислительных машин в деле технического прогресса нашей страны исключительно велико. Современные электронные цифровые вычислительные машины производят десятки тысяч арифметических и логических операций в секунду и способны в исключительно короткие сроки давать решения сложнейших математических и технических задач, немыслимые при ручном счете.

Огромное быстродействие вычислительных машин открывает новые широкие возможности для применения общих математических методов исследования в проблемах физики, механики, химии, астрономии, техники, экономики и многих других областей.

Принципиальное отличие от прежнего положения вещей состоит в том, что сложнейшие технические и экономические вопросы могут решаться в точной постановке, недоступной для малых вычислительных машин из-за чрезмерного объема работы. В частности, например, представляется по-новому организация процесса технического проектирования с отказом от грубо ориентировочных расчетов.

Исключительное значение имеют электронные машины для автоматического управления быстро движущимися объектами, например межпланетными ракетными снарядами. Велика также роль электронных вычислительных машин для развития самой математики. Машины используются для подсчета математических постоянных; для решений алгебраических, трансцендентных и дифференциальных уравнений; для решения сложнейших функциональных неравенств и т. п. Появились новые статистические методы машинного решения задач математической физики, стало возможным экспериментальное решение логических задач и многое другое.

Таким образом, создание электронных вычислительных машин знаменует решительный скачок по пути прогресса точных и технических наук нашего времени.

Все это сделало весьма актуальным усовершенствование и развитие численных и приближенных методов решения задач. Дело в том,

что машина способна выполнять очень большое, но конечное число операций. Поэтому точные предельные процессы решения задач, связанные с бесконечным числом операций, при работе на машине по необходимости должны быть заменены приближенными алгоритмами, содержащими лишь конечное число действий. Например, при машинном вычислении определенного интеграла последний обычно заменяется конечной интегральной суммой, вместо дифференциального уравнения рассматривается конечная система уравнений в конечных разностях и т. п. Кроме того, машина обладает конечной памятью и может оперировать с числами лишь конечной длины. Поэтому промежуточные результаты округляются, в результате чего даже точный метод с конечным числом действий становится приближенным.

В настоящее время разработка численных и приближенных методов решения задач в основном протекает в двух направлениях: с одной стороны, создаются более эффективные детерминированные способы решения задач, учитывающие специфические особенности счетных машин; с другой стороны, в практику успешно внедряются статистические недетерминированные методы, основанные на случайных испытаниях (метод Монте-Карло и другие) [1], [2].

Высокая производительность электронных машин существенным образом изменила подход к оценке того или иного вычислительного метода. Ценным оказывается тот метод, который является наиболее универсальным и который допускает простую реализацию на машинах. Напротив, метод, основанный на частных особенностях задачи или на искусстве вычислителя, оказывается теперь мало пригодным. В связи с этим произошла своеобразная переоценка ценностей: многие вычислительные методы, приводившие к громоздкому счету и считавшиеся раньше, при ручном счете, непрактичными, оказались сейчас вполне рабочими. В то же время чисто аналитические конструкции, ведущие к неудобным вычислительным алгоритмам, потеряли свою былую ценность. Вот почему сейчас большое распространение получили итеративные, разностные, вариационные, вероятностные и т. п. методы решения задач, допускающие удобные схемы счета и применимые к широкому кругу проблем.

При приближенном решении задач необходима оценка погрешности полученного результата. Здесь при вычислениях с большим числом шагов мы сталкиваемся с новой весьма важной проблемой — вопросом устойчивости вычислительной схемы. Может случиться, что неизбежные погрешности округлений быстро накапливаются (например, имеют показательный рост). Такая вычислительная схема неустойчива и непригодна для практики. Допустимо пользоваться лишь устойчивыми вычислительными схемами, когда погрешности округлений взаимно компенсируются и вызываемая ими ошибка результата остается малой для всего процесса вычислений.

Выбирая численный метод среди известных методов или разрабатывая новый, мы должны, естественно, учитывать специфику машины, на которой предполагается решать задачу.

Описание принципов устройства ЭВМ, а также технику решения на них математических задач читатель найдет в специальных руководствах (см., например, [3]—[10]).

ЛИТЕРАТУРА К ВВЕДЕНИЮ

[1] Математика, ее содержание, методы и значение, Изд. АН СССР, 1956, т. 2, гл. XIV.
[2] Бусленко Н. П., Голенко Д. И., Соболь И. М., Срагович В. Г., Шрейдер Ю. А., Метод статистических испытаний, СМБ, Физматгиз, 1962.
[3] Гутер Р. С., Овчинский Б. В., Резниковский П. Т., Программирование и вычислительная математика, «Наука», 1965.
[4] Криницкий Н. А., Миронов Г. А., Фролов Г. Д., Программирование, СМБ, Физматгиз, 1961.
[5] Гнеденко Б. В., Королюк В. С., Ющенко Е. Л., Элементы программирования, Физматгиз, 1961.
[6] Ричардс Р. К., Элементы и схемы цифровых вычислительных машин, ИЛ, 1961.
[7] Система автоматизации программирования, под ред. Трифонова Н. А. и Шура-Бура М. Р., Физматгиз, 1961.
[8] Люстерник Л. А., Абрамов А. А., Шестаков В. И., Шура-Бура М. Р., Решение математических задач на автоматических машинах, Изд. АН СССР, 1952.
[9] Михлин С. Г., Смолицкий Х. Л., Приближенные методы решения дифференциальных и интегральных уравнений, «Наука», 1965.
[10] Ланцош К., Практические методы прикладного анализа, Физматгиз, 1961.

ГЛАВА I

ПРИБЛИЖЕНИЕ ФУНКЦИЙ

§ 1. Постановка задачи о приближении функций

Пусть на некотором множестве задана система функций $\varphi_0(x)$, $\varphi_1(x), \ldots, \varphi_m(x), \ldots$, которые в дальнейшем будем считать достаточно гладкими (например, непрерывно дифференцируемыми) функциями. Назовем эту систему *основной*.

Функции вида

$$Q_m(x) = c_0\varphi_0(x) + c_1\varphi_1(x) + \ldots + c_m\varphi_m(x), \tag{1}$$

где $c_0, c_1, \ldots, c_m$ — постоянные коэффициенты, называются *обобщенными полиномами (обобщенными многочленами)* порядка m. В частности, если основная система состоит из целых неотрицательных степеней переменной x, т. е. $\varphi_0(x) = 1$, $\varphi_1(x) = x, \ldots, \varphi_m(x) = x^m, \ldots$, то

$$Q_m(x) = c_0 + c_1x + \ldots + c_mx^m$$

есть обычный полином степени m.

Если $\varphi_0(x) = 1$, $\varphi_1(x) = \cos x$, $\varphi_2(x) = \sin x, \ldots, \varphi_{2m-1}(x) = \cos mx$, $\varphi_{2m}(x) = \sin mx, \ldots$, то

$$Q_m(x) = a_0 + a_1\cos x + b_1\sin x + \ldots + a_m\cos mx + b_m\sin mx$$

называется *тригонометрическим полиномом* (или *тригонометрическим многочленом*) *порядка* m.

Задача о приближении функции ставится следующим образом: данную функцию $f(x)$ требуется заменить обобщенным полиномом $Q_m(x)$ заданного порядка m так, чтобы отклонение (в известном смысле) функции $f(x)$ от обобщенного полинома $Q_m(x)$ на указанном множестве $X = \{x\}$ было наименьшим. При этом полином $Q_m(x)$ в общем случае называется *аппроксимирующим*.

Если множество X состоит из отдельных точек $x_0, x_1, \ldots, x_n$, то приближение называется *точечным*. Если же X есть отрезок $a \leqslant x \leqslant b$, то приближение называется *интегральным*.

Для практики весьма важным является приближение функций обычными и тригонометрическими полиномами.

Что касается термина отклонение двух функций, то он в зависимости от обстоятельств понимается по-разному. При этом мы получаем различные типы задач теории приближений: интерполирование, среднеквадратичное приближение, равномерное приближение и т. п. В следующих параграфах остановимся подробно на некоторых из этих задач.

§ 2. Интерполирование функций

Будем считать данную функцию $f(x)$ и полином $Q_m(x)=a_0+a_1x+a_2x^2+\ldots+a_mx^m$ близкими, если они совпадают на заданной системе точек $x_0, x_1, \ldots, x_n$. Эти точки называются *узлами интерполирования* или *узлами интерполяции*. Таким образом, мы приходим к следующей задаче интерполирования: для данной функции $f(x)$ найти полином $Q_m(x)$ возможно низшей степени m, принимающий в заданных точках x_i $(i=0, 1, 2, \ldots, n;\ x_i \neq x_j$ при $i \neq j)$ те же значения, что и $f(x)$, т. е. такой, что

$$Q_m(x_i)=f(x_i) \qquad (i=0, 1, 2, \ldots, n).$$

Такой полином $Q_m(x)$ называется *интерполяционным*.

Как известно [1], существует единственный полином степени не выше n, принимающий в точках $x_0, x_1, \ldots, x_n$ заданные значения. Поэтому можно положить $m=n$. Коэффициенты a_i полинома $Q_n(x)$ можно определить из системы уравнений

$$\left.\begin{array}{l} a_0+a_1x_0+a_2x_0^2+\ldots+a_nx_0^n=y_0, \\ a_0+a_1x_1+a_2x_1^2+\ldots+a_nx_1^n=y_1, \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ a_0+a_1x_n+a_2x_n^2+\ldots+a_nx_n^n=y_n, \end{array}\right\} \qquad (1)$$

где $y_i=f(x_i)$ $(i=0, 1, 2, \ldots, n)$. Определитель этой системы есть так называемый *определитель Вандермонда* [1]

$$\Delta=\begin{vmatrix} 1 & x_0 & x_0^2 & \ldots & x_0^n \\ 1 & x_1 & x_1^2 & \ldots & x_1^n \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 1 & x_n & x_n^2 & \ldots & x_n^n \end{vmatrix}=\prod_{0\leqslant p<q\leqslant n}(x_q-x_p)\neq 0,$$

и, следовательно, система (1) имеет единственное решение.

Полином $L_n(x)$, коэффициенты которого определяются из системы (1), называется *интерполяционным полиномом Лагранжа* для функции $f(x)$ и может быть записан в явном виде [1], [2]:

$$L_n(x)=\sum_{i=0}^{n}\frac{(x-x_0)(x-x_1)\ldots(x-x_{i-1})(x-x_{i+1})\ldots(x-x_n)}{(x_i-x_0)(x_i-x_1)\ldots(x_i-x_{i-1})(x_i-x_{i+1})\ldots(x_i-x_n)}y_i. \qquad (2)$$

Для случая равноотстоящих узлов $x_i = x_0 + ih$ $(i = 0, 1, 2, \ldots, n)$ полином Лагранжа $L_n(x)$ может быть записан в виде интерполяционного полинома Ньютона [2]:

$$P_n(x) = y_0 + \frac{\Delta y_0}{h}(x - x_0) + \frac{\Delta^2 y_0}{2!\,h^2}(x - x_0)(x - x_1) + \ldots + \\ + \frac{\Delta^n y_0}{n!\,h^n}(x - x_0)\ldots(x - x_n), \qquad (3)$$

где

$$\begin{aligned} \Delta y_i &= \Delta f(x_i) = f(x_{i+1}) - f(x_i), \\ \Delta^2 y_i &= \Delta^2 f(x_i) = \Delta y_{i+1} - \Delta y_i, \\ &\ldots\ldots\ldots\ldots \\ \Delta^n y_i &= \Delta^n f(x_i) = \Delta^{n-1} y_{i+1} - \Delta^{n-1} y_i \\ &(i = 0, 1, 2, \ldots, n-1) \end{aligned}$$

— конечные разности различных порядков.

Полагая

$$\frac{x - x_0}{h} = q,$$

можно интерполяционный полином Ньютона (3) записать в следующем виде:

$$P_n(x) = y_0 + q\Delta y_0 + \frac{q(q-1)}{2!}\Delta^2 y_0 + \ldots + \frac{q(q-1)\ldots(q-n+1)}{n!}\Delta^n y_0. \qquad (3')$$

Если функция $y = f(x)$ задана таблично, причем известны ее значения $y_i = f(x_i)$ $(i = 0, 1, 2, \ldots, n)$ лишь в узлах $x_0, x_1, \ldots, x_n$, то интерполяционный полином Лагранжа $L_n(x)$ может быть использован для приближенного нахождения нетабличных значений функции. Эта операция носит название *интерполирования функции*. В случае равноотстоящих узлов для интерполирования функции можно воспользоваться интерполяционным полиномом Ньютона $P_n(x)$. В частности, при $n = 1$ имеем формулу *линейного интерполирования*

$$y \approx y_0 + \frac{y_1 - y_0}{h}(x - x_0), \qquad (4)$$

а при $n = 2$ — формулу *квадратичного интерполирования*

$$y \approx y_0 + \frac{y_1 - y_0}{h}(x - x_0) + \frac{y_2 - 2y_1 + y_0}{2h^2}(x - x_0)(x - x_1), \qquad (5)$$

где $h = \Delta x_i = x_{i+1} - x_i$ $(i = 0, 1, 2, \ldots)$ — шаг таблицы.

При определении значений функции y в точке x по формуле (3) в качестве x_0, вообще говоря, выгодно брать ближайший к x узел интерполирования; тогда, если шаг h достаточно мал, то первое слагаемое этой формулы будет главным членом, а остальные слагаемые будут носить характер поправок.

Если функция $y=f(x)$ непрерывно дифференцируема до $(n+1)$-го порядка включительно, то *остаточный член* ее интерполяционного полинома Лагранжа имеет вид [2]

$$R_n(x)=f(x)-L_n(x)=\frac{f^{(n+1)}(\xi)}{(n+1)!}(x-x_0)(x-x_1)\ldots(x-x_n), \quad (6)$$

где ξ — внутренняя точка минимального отрезка, содержащего узлы интерполирования $x_0, x_1, \ldots, x_n$ и точку x. В случае, если функция $f(x)$ задана таблично на системе равноотстоящих точек $x_i=x_0+ih$ $(i=0, 1, 2, \ldots, n)$, то при достаточно малом шаге h производную $f^{(n+1)}(\xi)$ приближенно можно заменить конечно-разностным отношением

$$f^{(n+1)}(\xi)\approx\frac{\Delta^{n+1}y_0}{h^{n+1}},$$

где

$$y_0=f(x_0).$$

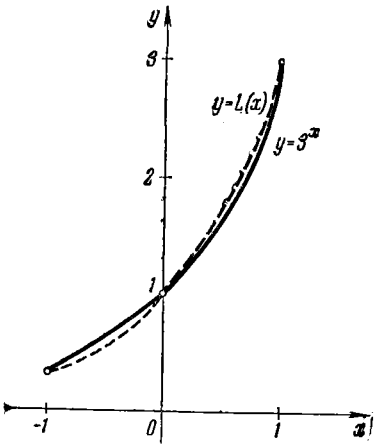

Рис. 1.

Пример 1. Построить интерполяционный полином $L(x)$, совпадающий с функцией $f(x)=3^x$ $(-1\leqslant x\leqslant 1)$ в точках $x_0=-1$, $x_1=0$, $x_2=1$.

Решение. Полагаем $L(x)=a_0+a_1x+a_2x^2$. Для определения коэффициентов a_0, a_1 и a_2 имеем систему

$$\left.\begin{aligned} a_0-a_1+a_2&=\frac{1}{3},\\ a_0\qquad\quad&=1,\\ a_0+a_1+a_2&=3.\end{aligned}\right\}$$

Отсюда $a_0=1$, $a_1=\frac{4}{3}$, $a_2=\frac{2}{3}$, и, следовательно (рис. 1),

$$3^x\approx 1+\frac{4}{3}x+\frac{2}{3}x^2 \quad \text{при} \quad -1\leqslant x\leqslant 1.$$

Например, $3^{\frac{1}{2}}=\sqrt{3}\approx 1+\frac{2}{3}+\frac{1}{6}\approx 1{,}8$.

Пример 2. Построить интерполяционный полином Лагранжа $L_4(x)$, совпадающий с функцией $y=2\cos\frac{\pi x}{4}$ в точках $x_0=-2$, $x_1=-\frac{4}{3}$, $x_2=0$, $x_3=\frac{4}{3}$, $x_4=2$.

Решение. Вычисляем значение данной функции в указанных узлах и составляем таблицу:

x	-2	$-4/3$	0	4/3	2
y	0	1	2	1	0

Подставляя эти значения в формулу (2), находим

$$L_4(x)=1\cdot\frac{(x+2)(x-0)\left(x-\frac{4}{3}\right)(x-2)}{\left(-\frac{4}{3}+2\right)\left(-\frac{4}{3}-0\right)\left(-\frac{4}{3}-\frac{4}{3}\right)\left(-\frac{4}{3}-2\right)}+$$

$$+2\cdot\frac{(x+2)\left(x+\frac{4}{3}\right)\left(x-\frac{4}{3}\right)(x-2)}{(0+2)\left(0+\frac{4}{3}\right)\left(0-\frac{4}{3}\right)(0-2)}+$$

$$+1\cdot\frac{(x+2)\left(x+\frac{4}{3}\right)(x-0)(x-2)}{\left(\frac{4}{3}+2\right)\left(\frac{4}{3}+\frac{4}{3}\right)\left(\frac{4}{3}-0\right)\left(\frac{4}{3}-2\right)}=\frac{9x^4-196x^2+640}{320}=$$

$$=0{,}0281x^4-0{,}6125x^2+2.$$

Заметим, что $L_4(1)=1{,}4156$, а $y(1)=2\cos\frac{\pi}{4}=1{,}4142$, т. е. $L_4(1)-y(1)=$ $=0{,}0014$.

Пример 3. Используя интерполяционный полином Ньютона, вычислить $f(0{,}14)$, где функция $y=f(x)$ задана таблицей

x	0	0,1	0,2	0,3	0,4	0,5
y	0	0,1002	0,2013	0,8045	0,4108	0,5211

Решение. Составляем таблицу конечных разностей:

x	y	Δy	$\Delta^2 y$	$\Delta^3 y$	$\Delta^4 y$	$\Delta^5 y$
0,	0	0,1002	0,0009	0,0012	—0,0002	0,0001
0,1	0,1002	0,1011	0,0021	0,0010	—0,0001	
0,2	0,2013	0,1032	0,0031	0,0009		
0,3	0,3045	0,1063	0,0040			
0,4	0,4108	0,1103				
0,5	0,5211					

Для вычисления $f(0{,}14)$ воспользуемся интерполяционным полиномом Ньютона (3), полагая $x_0=0{,}1$ и $h=0{,}1$; тогда

$$q=\frac{0{,}14-0{,}1}{0{,}1}=0{,}4$$

и

$$f(0{,}14)\approx 0{,}1002+0{,}1011\cdot 0{,}4\cdot\frac{0{,}0021}{2}\cdot 0{,}4\cdot(-0{,}6)+\frac{0{,}1010}{3!}0{,}4\cdot(-0{,}6)(-1{,}6)\approx$$

$$\approx 0{,}1405.$$

При этом погрешность приближения

$$R_4 = |f(0{,}14) - P_3(0{,}14)| < \frac{0{,}0001}{4!} 0{,}4 \cdot 0{,}6 \cdot 1{,}6 \cdot 2{,}6 = 4{,}16 \cdot 10^{-6} < \frac{1}{2} \cdot 10^{-5}.$$

При интерполировании функции обобщенными полиномами

$$Q_n(x) = c_0\varphi_0(x) + c_1\varphi_1(x) + \ldots + c_n\varphi_n(x)$$

коэффициенты $c_0, c_1, \ldots, c_n$ находим из системы уравнений

$$\left.\begin{array}{l} c_0\varphi_0(x_0) + c_1\varphi_1(x_0) + \ldots + c_n\varphi_n(x_0) = f(x_0), \\ c_0\varphi_0(x_1) + c_1\varphi_1(x_1) + \ldots + c_n\varphi_n(x_1) = f(x_1), \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ c_0\varphi_0(x_n) + c_1\varphi_1(x_n) + \ldots + c_n\varphi_n(x_n) = f(x_n). \end{array}\right\} \quad (7)$$

Если функции $\varphi_0(x), \varphi_1(x), \ldots, \varphi_n(x)$ линейно независимы, то определитель системы

$$\Delta = \begin{vmatrix} \varphi_0(x_0) & \varphi_1(x_0) & \ldots & \varphi_n(x_0) \\ \varphi_0(x_1) & \varphi_1(x_1) & \ldots & \varphi_n(x_1) \\ \cdots & \cdots & \cdots & \cdots \\ \varphi_0(x_n) & \varphi_1(x_n) & \ldots & \varphi_n(x_n) \end{vmatrix} \neq 0,$$

и, следовательно, система (7) имеет единственное решение.

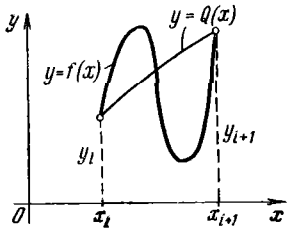

Рис. 2.

Важным случаем общего интерполирования функции является тригонометрическое интерполирование (см. § 3).

Интерполирование не всегда дает удовлетворительное решение задачи о приближении функции с заданной точностью на данном промежутке, так как совпадение функции $f(x)$ с полиномом $Q(x)$ даже в близких точках x_i и x_{i+1} не гарантирует малость величины $|f(x) - Q(x)|$ на отрезке $[x_i, x_{i+1}]$ (рис. 2).

§ 3. Интерполирование периодических функций с помощью тригонометрических полиномов

Пусть $f(x)$ периодическая и задана на оси $-\infty < x < +\infty$ функция. Путем линейной замены независимой переменной период функции можно сделать равным $T = 2\pi$. В этом случае функцию $f(x)$ целесообразно интерполировать тригонометрическим полиномом

$$Q_n(x) = a_0 + \sum_{k=1}^{n} (a_k \cos kx + b_k \sin kx) \quad (1)$$

так, чтобы

$$Q_n(x_i) = f(x_i) \qquad (i = 0, 1, 2, \ldots, 2n), \quad (2)$$

где $0 \leqslant x_0 < x_1 < \ldots < x_{2n} < 2\pi$ — точки промежутка $[0, 2\pi)$. Полином $Q_n(x)$ будем называть *тригонометрическим полиномом порядка n.*

Пусть $y_i = f(x_i)$ $(i = 0, 1, 2, \ldots, 2n)$. Требуется подобрать коэффициенты полинома (1) так, чтобы выполнялись следующие равенства:

$$\left.\begin{aligned} y_0 &= a_0 + \sum_{k=1}^{n} (a_k \cos kx_0 + b_k \sin kx_0), \\ y_1 &= a_0 + \sum_{k=1}^{n} (a_k \cos kx_1 + b_k \sin kx_1), \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ y_{2n} &= a_0 + \sum_{k=1}^{n} (a_k \cos kx_{2n} + b_k \sin kx_{2n}). \end{aligned}\right\} \tag{3}$$

Мы получили систему $2n+1$ линейных уравнений с $2n+1$ неизвестными $a_0, a_1, b_1, \ldots, a_n, b_n$.

Как известно [1], определитель этой системы

$$\Delta = \begin{vmatrix} 1 & \cos x_0 & \sin x_0 & \ldots & \cos nx_0 & \sin nx_0 \\ 1 & \cos x_1 & \sin x_1 & \ldots & \cos nx_1 & \sin nx_1 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 1 & \cos x_{2n} & \sin x_{2n} & \ldots & \cos nx_{2n} & \sin nx_{2n} \end{vmatrix} = 2^{n^2} \prod_{0 \leqslant p < q \leqslant 2n} \sin \frac{x_q - x_p}{2}$$

и, следовательно, отличен от нуля для нашей системы точек, для которой $0 < x_q - x_p < 2\pi$. Поэтому данная интерполяционная задача имеет решение, и притом единственное.

Таким образом, справедливо следующее утверждение:

Теорема единственности. *Тригонометрический полином* (1) *порядка n однозначно определяется своими значениями в $2n+1$ различных точках, расположенных на промежутке* $[0, 2\pi)$.

Для построения полинома $Q_n(x)$ возьмем произвольную точку x, не совпадающую с узлами x_i $(i = 0, 1, 2, \ldots, 2n)$, и для этих $2n+2$ точек составим систему из $2n+2$ уравнений

$$\left.\begin{aligned} Q_n(x) - a_0 - \sum_{k=1}^{n} (a_k \cos kx + b_k \sin kx) &= 0, \\ f(x_0) - a_0 - \sum_{k=1}^{n} (a_k \cos kx_0 + b_k \sin kx_0) &= 0, \\ f(x_1) - a_0 - \sum_{k=1}^{n} (a_k \cos kx_1 + b_k \sin kx_1) &= 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ f(x_{2n}) - a_0 - \sum_{k=1}^{n} (a_k \cos kx_{2n} + b_k \sin kx_{2n}) &= 0, \end{aligned}\right\}$$

линейных и однородных относительно коэффициентов $c=1$ при $Q_n(x)$, $f(x_k)$ и a_k, b_k $(k=0, 1, \ldots, 2n)$. Эта система, очевидно, имеет ненулевое решение и поэтому ее определитель равен нулю, т. е.

$$\begin{vmatrix} Q_n(x) & 1 & \cos x & \sin x & \ldots & \cos nx & \sin nx \\ f(x_0) & 1 & \cos x_0 & \sin x_0 & \ldots & \cos nx_0 & \sin nx_0 \\ f(x_1) & 1 & \cos x_1 & \sin x_1 & \ldots & \cos nx_1 & \sin nx_1 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ f(x_{2n}) & 1 & \cos x_{2n} & \sin x_{2n} & \ldots & \cos nx_{2n} & \sin nx_{2n} \end{vmatrix} = 0.$$

Раскладывая этот определитель по элементам первого столбца и разрешая полученное уравнение относительно $Q_n(x)$, после упрощений [1] находим искомый тригонометрический полином

$$Q_n(x) = \sum_{i=0}^{2n} f(x_i) \frac{\sin\frac{x-x_0}{2} \ldots \sin\frac{x-x_{i-1}}{2} \sin\frac{x-x_{i+1}}{2} \ldots \sin\frac{x-x_{2n}}{2}}{\sin\frac{x_i-x_0}{2} \ldots \sin\frac{x_i-x_{i-1}}{2} \sin\frac{x_i-x_{i+1}}{2} \ldots \sin\frac{x_i-x_{2n}}{2}}. \tag{4}$$

Легко непосредственно убедиться, что правая часть формулы (4) есть тригонометрический полином порядка n, удовлетворяющий условиям (2). Действительно, используя формулы

$$\sin\alpha \sin\beta = \frac{1}{2}[\cos(\alpha-\beta) - \cos(\alpha+\beta)],$$

$$\cos\alpha \cos\beta = \frac{1}{2}[\cos(\alpha-\beta) + \cos(\alpha+\beta)],$$

$$\sin\alpha \cos\beta = \frac{1}{2}[\sin(\alpha-\beta) + \sin(\alpha+\beta)]$$

и учитывая, что числитель каждого слагаемого формулы (4) содержит $2n$ сомножителей, получим, что сумма $Q_n(x)$ есть тригонометрический полином порядка n. Кроме того, очевидно, имеем

$$Q_n(x_j) = f(x_j) \qquad (j=0, 1, \ldots, 2n).$$

В силу теоремы единственности найденный полином искомый.

По своей структуре полином $Q_n(x)$ напоминает интерполяционный полином Лагранжа $L_n(x)$.

Пример 1. Построить тригонометрический полином второго порядка, принимающий в точках $x_0=0$, $x_1=\pi/4$, $x_2=\pi/2$, $x_3=\pi$, $x_4=3\pi/2$ соответственно значения $y_0=2$, $y_1=1$, $y_2=0$, $y_3=4$, $y_4=-2$.

Решение. Согласно формуле (4) искомый полином будет следующим:

$$Q_2(x)=2\cdot\frac{\sin\frac{x-\pi/4}{2}\sin\frac{x-\pi/2}{2}\sin\frac{x-\pi}{2}\sin\frac{x-3\pi/2}{2}}{\sin\frac{0-\pi/4}{2}\sin\frac{0-\pi/2}{2}\sin\frac{0-\pi}{2}\sin\frac{0-3\pi/2}{2}}+$$

$$+1\cdot\frac{\sin\frac{x-0}{2}\sin\frac{x-\pi/2}{2}\sin\frac{x-\pi}{2}\sin\frac{x-3\pi/2}{2}}{\sin\frac{\pi/4-0}{2}\sin\frac{\pi/4-\pi/2}{2}\sin\frac{\pi/4-\pi}{2}\sin\frac{\pi/4-3\pi/2}{2}}+$$

$$+4\cdot\frac{\sin\frac{x-0}{2}\sin\frac{x-\pi/4}{2}\sin\frac{x-\pi/2}{2}\sin\frac{x-3\pi/2}{2}}{\sin\frac{\pi-0}{2}\sin\frac{\pi-\pi/4}{2}\sin\frac{\pi-\pi/2}{2}\sin\frac{\pi-3\pi/2}{2}}-$$

$$-2\cdot\frac{\sin\frac{x-0}{2}\sin\frac{x-\pi/4}{2}\sin\frac{x-\pi/2}{2}\sin\frac{x-\pi}{2}}{\sin\frac{3\pi/2-0}{2}\sin\frac{3\pi/2-\pi/4}{2}\sin\frac{3\pi/2-\pi/2}{2}\sin\frac{3\pi/2-\pi}{2}}.$$

После упрощений окончательно получаем $Q_2(x)=1+\sin x-\cos x+2\cos 2x$.

Замечание. Если $f(x)$, периодическая функция периода T, задана в точках $x_0, x_1, \ldots, x_{2n}$ промежутка $[a, a+T)$ (a — любое число), то с помощью линейного преобразования $t=\frac{2\pi}{T}(x-a)$ получаем

$$f(x)=f\left(a+\frac{2\pi}{T}t\right)\equiv\varphi(t),$$

где $\varphi(t)$ имеет период 2π и задана в точках

$$t_i=\frac{2\pi}{T}(x_i-a)\in[0, 2\pi) \qquad (i=0, 1, 2, \ldots, 2n).$$

Таким образом, мы приходим к разобранному выше случаю.

Если точки интерполирования x_p $(p=0, 1, \ldots, 2n)$ равноотстоящие, т. е. $x_p=\frac{2p\pi}{2n+1}$ $(p=0, 1, \ldots, 2n)$, то интерполяционному тригонометрическому полиному (4) можно придать более простой вид:

$$Q_n(x)=\frac{1}{2n+1}\sum_{p=0}^{2n}f(x_p)\frac{\sin(2n+1)\frac{x-x_p}{2}}{\sin\frac{x-x_p}{2}}. \tag{5}$$

Непосредственно легко убедиться, что

$$Q_n(x_q)=f(x_q) \qquad (q=0, 1, \ldots, 2n).$$

Действительно, при $p\neq q$, учитывая, что $x_q-x_p=\frac{2\pi(q-p)}{2n+1}$, имеем

$$\sin(2n+1)\frac{x_q-x_p}{2}=\sin\pi(q-p)=0,$$

и, следовательно, соответствующие слагаемые суммы (5) равны нулю. При $p=q$, используя предельный переход, получаем

$$\lim_{x\to x_q}\frac{\sin(2n+1)\dfrac{x-x_q}{2}}{\sin\dfrac{x-x_q}{2}}=2n+1.$$

Поэтому $Q_n(x_q)=f(x_q)$.

Пример 2. Функция $y=f(x)$ задана таблицей:

x	0	$\frac{2\pi}{7}$	$\frac{4\pi}{7}$	$\frac{6\pi}{7}$	$\frac{8\pi}{7}$	$\frac{10\pi}{7}$	$\frac{12\pi}{7}$	2π
y	0	0,5029	0,8600	1,1058	1,2058	1,1227	0,7855	0

Используя тригонометрическое интерполирование, найти $f(\pi/7)$.

Решение. Здесь $2n+1=7$, $x=\pi/7$. Согласно формуле (5) находим

$$f\left(\frac{\pi}{7}\right)\approx Q_3(x)=\frac{1}{7}\left[0\cdot\frac{\sin\frac{\pi}{2}}{\sin\frac{\pi}{14}}+0{,}5029\cdot\frac{\sin\left(-\frac{\pi}{2}\right)}{\sin\left(-\frac{\pi}{14}\right)}+0{,}8600\cdot\frac{\sin\left(-\frac{3\pi}{2}\right)}{\sin\left(-\frac{3\pi}{14}\right)}+\right.$$

$$+1{,}058\cdot\frac{\sin\left(-\frac{5\pi}{2}\right)}{\sin\left(-\frac{5\pi}{14}\right)}+1{,}2053\cdot\frac{\sin\left(-\frac{7\pi}{2}\right)}{\sin\left(-\frac{7\pi}{14}\right)}+1{,}1227\cdot\frac{\sin\left(-\frac{9\pi}{2}\right)}{\sin\left(-\frac{9\pi}{14}\right)}+$$

$$\left.+0{,}7855\cdot\frac{\sin\left(-\frac{11\pi}{2}\right)}{\sin\left(-\frac{11\pi}{14}\right)}\right]=$$

$$=\frac{1}{7}\left(0+\frac{0{,}5029}{0{,}2225}-\frac{0{,}8600}{0{,}6235}+\frac{1{,}058}{0{,}9010}-1{,}2053+\frac{1{,}1227}{0{,}6235}-\frac{0{,}7855}{0{,}9010}\right)=0{,}2617.$$

Заметим, что приведенная таблица составлена для функции $f(x)=\left(2-\frac{x}{\pi}\right)\operatorname{sh}\frac{x}{\pi}$. Так как $f(\pi/7)=0{,}2618$, то абсолютная погрешность найденного значения не превышает 10^{-4}.

§ 4. Точечное квадратичное аппроксимирование функций

На практике часто бывает, что заданный порядок m приближающего полинома $Q_m(x)$ значительно меньше числа узлов n. В этом случае интерполирование, вообще говоря, становится невозможным и приходится прибегать к иным приемам построения приближающего полинома для данной функции. Обычно здесь используют *точечный*

способ наименьших квадратов. Согласно этому способу за меру отклонения полинома

$$Q_m(x) = a_0 + a_1 x + \ldots + a_m x^m \tag{1}$$

от данной функции $f(x)$ на множестве точек $x_0, x_1, \ldots, x_n$ принимают величину

$$S_m = \sum_{i=0}^{n} [Q_m(x_i) - f(x_i)]^2, \tag{2}$$

равную сумме квадратов отклонений полинома $Q_m(x)$ от функции $f(x)$ на заданной системе точек.

Очевидно, что S_m есть функция коэффициентов $c_0, c_1, \ldots, c_m$. Эти коэффициенты надо подобрать так, чтобы величина S_m была наименьшей. Полученный полином называется *аппроксимирующим* для данной функции, а процесс построения этого полинома — *точечной квадратичной аппроксимацией* или *точечным квадратичным аппроксимированием* функции.

Для решения задачи точечного квадратичного аппроксимирования воспользуемся общим приемом дифференциального исчисления. Найдем частные производные от величины

$$S_m = \sum_{i=0}^{n} (a_0 + a_1 x_i + a_2 x_i^2 + \ldots + a_m x_i^m - y_i)^2,$$

где $y_i = f(x_i)$, по всем переменным $a_0, a_1, \ldots, a_m$. Приравнивая эти частные производные нулю, получим для определения неизвестных $a_0, a_1, \ldots, a_m$ систему $m+1$ уравнений с $m+1$ неизвестными:

$$\left.\begin{aligned}
&\frac{1}{2}\frac{\partial S_m}{\partial a_0} = \sum_{i=0}^{n} (a_0 + a_1 x_i + a_2 x_i^2 + \ldots + a_m x_i^m - y_i)\, 1 = 0,\\
&\frac{1}{2}\frac{\partial S_m}{\partial a_1} = \sum_{i=0}^{n} (a_0 + a_1 x_i + a_2 x_i^2 + \ldots + a_m x_i^m - y_i)\, x_i = 0,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&\frac{1}{2}\frac{\partial S_m}{\partial a_m} = \sum_{i=0}^{n} (a_0 + a_1 x_i + a_2 x_i^2 + \ldots + a_m x_i^m - y_i)\, x_i^m = 0.
\end{aligned}\right\} \tag{3}$$

Введем обозначения:

$$s_k = x_0^k + x_1^k + \ldots + x_n^k \qquad (k = 0, 1, 2, \ldots),$$
$$t_k = x_0^k y_0 + x_1^k y_1 + \ldots + x_n^k y_n \qquad (k = 0, 1, 2, \ldots).$$

Преобразуя систему (3) и используя введенные обозначения, будем иметь

$$\left.\begin{aligned} a_0s_0+a_1s_1+a_2s_2+\ldots+a_ms_m&=t_0,\\ a_0s_1+a_1s_2+a_2s_3+\ldots+a_ms_{m+1}&=t_1,\\ a_0s_2+a_1s_3+a_2s_4+\ldots+a_ms_{m+2}&=t_2,\\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\ a_0s_m+a_1s_{m+1}+a_2s_{m+2}+\ldots+a_ms_{2m}&=t_m, \end{aligned}\right\} \qquad (4)$$

где $s_0=n+1$.

Можно доказать, что если среди точек $x_0, x_1, \ldots, x_n$ нет совпадающих и $m\leqslant n$, то определитель системы (4) отличен от нуля и, следовательно, эта система имеет единственное решение $a_0=a_0^*$, $a_1=a_1^*$, ..., $a_m=a_m^*$ [3]. Полином (1) с такими коэффициентами будет обладать минимальным квадратичным отклонением $S_{\min}$.

Если $m=n$, то аппроксимирующий полином $Q_m(x)$ совпадает с полиномом Лагранжа для системы точек $x_0, x_1, \ldots, x_m$, причем $S_{\min}=0$.

Таким образом, аппроксимирование функций представляет собой более общий процесс, чем интерполирование. При работе на счетно-электронных машинах для решения линейной системы (4) выгодно использовать итерационные методы [2]. В частности, так как матрица системы (4) положительно определенная, то для этой системы будет сходящимся процесс Зейделя [2].

Для составления системы (3) рекомендуется схема способа наименьших квадратов, приведенная в таблице 1, где принято $m=2$.

Таблица 1

Схема способа наименьших квадратов

x^0	x	x^2	x^3	x^4	y	xy	x^2y
1	x_0	x_0^2	x_0^3	x_0^4	y_0	x_0y_0	$x_0^2y_0$
1	x_1	x_1^2	x_1^3	x_1^4	y_1	x_1y_1	$x_1^2y_1$
1	x_2	x_2^2	x_2^3	x_2^4	y_2	x_2y_2	$x_2^2y_2$
1	x_3	x_3^2	x_3^3	x_3^4	y_3	x_3y_3	$x_3^2y_3$
1	x_4	x_4^2	x_4^3	x_4^4	y_4	x_4y_4	$x_4^2y_4$
s_0	s_1	s_2	s_3	s_4	t_0	t_1	t_2

Пример 1. Подобрать аппроксимирующий полином второй степени $y=a_0+a_1x+a_2x^2$ для данных [3]:

x	0,78	1,56	2,34	3,12	3,81
y	2,50	1,20	1,12	2,25	4,28

Решение. Вычисления, которые нам нужно произвести, расположим по схеме (для $m=2$, $n=4$), приведенной в таблице 1.

Для данного примера получаем таблицу 2 (вычисления проводятся с тремя десятичными знаками).

Таблица 2

Вычисления по способу наименьших квадратов

x^0	x	x^2	x^3	x^4	y	xy	x^2y
1	0,78	0,608	0,475	0,370	2,50	1,950	1,520
1	1,56	2,434	3,796	5,922	1,20	1,872	2,921
1	2,34	5,476	12,813	29,982	1,12	2,621	6,133
1	3,12	9,734	30,371	94,759	2,25	7,020	21,902
1	3,81	14,516	55,306	210,717	4,28	16,307	62,128
5	11,61	32,768	102,761	341,750	11,35	29,770	94,604

Отсюда система для определения коэффициентов a_0, a_1, a_2, имеет вид

$$\left.\begin{aligned} 5a_0+11{,}61a_1+32{,}768a_2&=11{,}350,\\ 11{,}61a_0+32{,}768a_1+102{,}761a_2&=29{,}770,\\ 32{,}768a_0+102{,}761a_1+341{,}750a_2&=94{,}604.\end{aligned}\right\}\qquad(5)$$

Решив систему (5), будем иметь: $a_0=5{,}045$; $a_1=-4{,}043$; $a_2=1{,}009$.

Следовательно, искомый полином есть

$$y=5{,}045-4{,}043x+1{,}009x^2.\qquad(6)$$

Сравним исходные значения для y с соответствующими значениями $\bar{y}$, полученными из приближенной формулы (6). Соответствующие результаты приведены в таблице 3.

Таблица 3

Погрешности вычисления по способу наименьших квадратов

x	y	$\bar{y}$	$\bar{y}-y$
0,78	2,50	2,505	0,005
1,56	1,20	1,194	—0,006
2,34	1,12	1,110	—0,010
3,12	2,25	2,252	0,002
3,81	4,28	4,288	0,008

В общем случае, когда аппроксимирующий полином для данной функции $f(x)$ является обобщенным:

$$Q_m(x) = c_0\varphi_0(x) + c_1\varphi_1(x) + \ldots + c_m\varphi_m(x),$$

для нахождения его коэффициентов $c_0, c_1, \ldots, c_m$ по способу наименьших квадратов приходится минимизировать сумму квадратов

$$S_m = \sum_{i=0}^{n} [c_0\varphi_0(x_i) + c_1\varphi_1(x_i) + \ldots + c_m\varphi_m(x_i) - f(x_i)]^2,$$

где $x_0, x_1, \ldots, x_n$ — заданная система точек. Используя необходимые условия экстремума функции, получаем систему уравнений

$$\left.\begin{aligned}
\frac{1}{2}\frac{\partial S_m}{\partial c_0} &= \sum_{i=0}^{n} [c_0\varphi_0(x_i) + c_1\varphi_1(x_i) + \ldots + c_m\varphi_m(x_i) - \\
&\qquad - f(x_i)]\,\varphi_0(x_i) = 0, \\
\frac{1}{2}\frac{\partial S_m}{\partial c_1} &= \sum_{i=0}^{n} [c_0\varphi_0(x_i) + c_1\varphi_1(x_i) + \ldots + c_m\varphi_m(x_i) - \\
&\qquad - f(x_i)]\,\varphi_1(x_i) = 0, \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
\frac{1}{2}\frac{\partial S_m}{\partial c_m} &= \sum_{i=0}^{n} [c_0\varphi_0(x_i) + c_1\varphi_1(x_i) + \ldots + c_m\varphi_m(x_i) - \\
&\qquad - f(x_i)]\,\varphi_m(x_i) = 0.
\end{aligned}\right\} \qquad (7)$$

Вводя сокращенные обозначения

$$(\varphi, \psi) = \sum_{i=0}^{n} \varphi(x_i)\,\psi(x_i),$$

систему (7) приводим к виду

$$\left.\begin{aligned}
&c_0(\varphi_0, \varphi_0) + c_1(\varphi_1, \varphi_0) + \ldots + c_m(\varphi_m, \varphi_0) = (f, \varphi_0), \\
&c_0(\varphi_0, \varphi_1) + c_1(\varphi_1, \varphi_1) + \ldots + c_m(\varphi_m, \varphi_1) = (f, \varphi_1), \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
&c_0(\varphi_0, \varphi_m) + c_1(\varphi_1, \varphi_m) + \ldots + c_m(\varphi_m, \varphi_m) = (f, \varphi_m).
\end{aligned}\right\} \qquad (8)$$

Из этой системы определяются коэффициенты $c_0, c_1, \ldots, c_m$.

Пример 2. Функцию

$$f(x) = 1 - \frac{x^2}{\pi^2}$$

по способу наименьших квадратов аппроксимировать на промежутке $[0, \pi]$ тригонометрическим полиномом

$$Q(x) = c_0 + c_1 \cos x$$

для системы точек $x_0 = 0$, $x_1 = \pi/3$, $x_2 = \pi/2$, $x_3 = 2\pi/3$, $x_4 = \pi$.

Решение. Полагая

$$\varphi_0(x)=1, \quad \varphi_1(x)=\cos x,$$

составляем таблицу 4 и таблицу 4а.

Таблица 4

Значения основных функций в узлах аппроксимирования

x	$\varphi_0(x)$	$\varphi_1(x)$	$\varphi_0^2(x)$	$\varphi_0(x)\varphi_1(x)$	$\varphi_1^2(x)$
0	1	1	1	1	1
$\pi/3$	1	1/2	1	1/2	1/4
$\pi/2$	1	0	1	0	0
$2\pi/3$	1	−1/2	1	−1/2	1/4
π	1	−1	1	−1	1
Σ			5	0	2,5

Таблица 4а

Вспомогательные вычисления для нормальной системы способа наименьших квадратов

x	$f(x)$	$f(x)\varphi_0(x)$	$f(x)\varphi_1(x)$
0	1	1	1
$\pi/3$	8/9	8/9	4/9
$\pi/2$	3/4	3/4	0
$2\pi/3$	5/9	5/9	−5/18
π	0	0	0
Σ		115/36	7/6

Отсюда находим:

$$(\varphi_0, \varphi_0)=5, \qquad (f, \varphi_0)=115/36,$$
$$(\varphi_0, \varphi_1)=0, \qquad (f, \varphi_1)=7/6.$$
$$(\varphi_1, \varphi_1)=2{,}5,$$

Следовательно, система для определения коэффициентов c_0 и c_1 имеет вид

$$5c_0=\frac{115}{36}, \quad 2{,}5c_1=\frac{7}{6},$$

откуда $c_0=\frac{23}{36}\approx 0{,}64$; $c_1=\frac{7}{15}\approx 0{,}47$. Таким образом,

$$Q(x)=0{,}64+0{,}47\cos x.$$

§ 5. Функции, ортогональные на точечном множестве

Если степень аппроксимирующего полинома сравнительно велика, то вычисления по способу наименьших квадратов становятся громоздкими. В этом случае иногда выгодно использовать новый метод построения аппроксимирующего полинома, основанный на понятии ортогональных функций.

Определение 1. Функции $\varphi(x)$ и $\psi(x)$ называются *ортогональными на множестве точек* $X=\{x_0, x_1, x_2, \ldots, x_n\}$, если

$$\sum_{i=0}^{n} \varphi(x_i)\,\psi(x_i)=0.$$

Например, функции $\varphi(x)=3x^2-15x+10$ и $\psi(x)=2x+5$ ортогональны на системе точек $x_i=i$ $(i=0, 1, 2, 3, 4, 5)$. В самом деле, так как

$$\varphi(0)=10,\ \varphi(1)=-2,\ \varphi(2)=-8,\ \varphi(3)=-8,\ \varphi(4)=-2,$$
$$\varphi(5)=10,$$
$$\psi(0)=5,\ \psi(1)=-3,\ \psi(2)=-1,\ \psi(3)=1,\ \psi(4)=3,\ \psi(5)=5,$$

то

$$\sum_{i=0}^{5} \varphi(x_i)\,\psi(x_i)=0.$$

Система функций $\{\varphi_k(x)\}$ называется *ортогональной на данном множестве* X, если функции системы попарно ортогональны между собой на этом множестве X.

Очевидно, функция $\varphi(x)$ обращающаяся в нуль в точках x_0, $x_1, \ldots, x_n$, ортогональна на этом множестве точек к любой другой функции. Поэтому всюду в дальнейшем будем предполагать, что

$$\sum_{i=0}^{n} \varphi^2(x_i)>0,$$

т. е. не все точки x_i $(i=0, 1, 2, \ldots, n)$ являются нулями рассматриваемых функций $\varphi(x)$.

Определение 2. Назовем величину

$$\sqrt{\sum_{i=0}^{n} \varphi^2(x_i)}=\|\varphi\|_X$$

нормой функции $\varphi(x)$ *на множестве* X.

Если система $\{\varphi_k(x)\}$ ортогональна на множестве X и для всех k выполнено равенство

$$\|\varphi_k(x)\|=1,$$

то такая система функций называется *ортонормированной*.

Легко видеть, что если система функций $\{\varphi_k(x)\}$ ортогональна на множестве X и не содержит функций с нулевой нормой, то система функций $\{\varphi_k(x)/\|\varphi_k\|_X\}$ — ортонормированная на X.

Определение 3. Функции $f_k(x)$ $(k=0, 1, 2, \ldots, m)$ называются *линейно независимыми на множестве* $X=\{x_0, x_1, x_2, \ldots, x_n\}$, если они определены на этом множестве и из равенств $\alpha_0 f_0(x_i)+\alpha_1 f_1(x_i)+\ldots+\alpha_m f_m(x_i)=0$ $(i=0, 1, 2, \ldots, n)$ следует, что все постоянные $\alpha_k=0$ $(k=0, 1, 2, \ldots, m)$. В противном случае функции $f_k(x)$ называются *линейно зависимыми на* X.

Лемма. *Функции* $\varphi_k(x)$ $(k=0, 1, 2, \ldots, m)$, *ортогональные на множестве* $X=\{x_0, x_1, x_2, \ldots, x_n\}$ *и имеющие ненулевые нормы, линейно независимы на этом множестве.*

Доказательство. Пусть

$$\alpha_0\varphi_0(x_i)+\alpha_1\varphi_1(x_i)+\alpha_2\varphi_2(x_i)+\ldots+\alpha_m\varphi_m(x_i)=0 \qquad (1)$$
$$(i=0, 1, 2, \ldots, n).$$

Умножая обе части равенства (1) на $\varphi_0(x_i)$ и суммируя по всем x_i, получим

$$\alpha_0\sum_{i=0}^{n}\varphi_0^2(x_i)+\alpha_1\sum_{i=0}^{n}\varphi_1(x_i)\varphi_0(x_i)+\ldots$$
$$\ldots+\alpha_m\sum_{i=0}^{n}\varphi_m(x_i)\varphi_0(x_i)=0. \qquad (2)$$

В силу ортогональности функций $\varphi_k(x)$ $(k=0, 1, \ldots, m)$ на множестве X все суммы в равенстве (2), начиная со второй, равны нулю. Поэтому

$$\alpha_0\sum_{i=0}^{n}\varphi_0^2(x_i)=0. \qquad (3)$$

Так как мы предполагаем

$$\|\varphi_0\|_X^2=\sum_{i=0}^{n}\varphi_0^2(x_i)>0,$$

то из равенства (3) имеем $\alpha_0=0$.

Аналогично доказывается, что $\alpha_k=0$ $(k=0, 1, 2, \ldots, n)$.

Таким образом, функции $\{\varphi_k(x)\}$ линейно независимы на множестве X.

Замечание. Понятие линейной независимости функций легко переносится на случай системы функций $\{f_k(x)\}$ $(k=0, 1, 2, \ldots, m)$, определенных на конечном или бесконечном промежутке $a<x<b$. А именно, функции $\{f_k(x)\}$ *линейно независимы на* (a, b), если тождество

$$\sum_{k=0}^{m}\alpha_k f_k(x)=0 \text{ при } x\in(a, b)$$

имеет место лишь тогда, когда $\alpha_0=\alpha_1=\alpha_2=\ldots=\alpha_m=0$. Очевидно, если функции $f_k(x)$ $(k=0, 1, 2, \ldots, m)$ линейно независимы на некотором точечном множестве $X\subset(a, b)$, то они будут также линейно независимы на промежутке (a, b).

Рассмотрим систему полиномов

$$P_0(x),\ P_1(x),\ \ldots,\ P_m(x), \tag{4}$$

ортогональных на данном множестве $X=\{x_0, x_1, \ldots, x_n\}$, т. е.

$$\sum_{i=0}^{n} P_j(x_i)P_k(x_i)=0 \quad \text{при } j\neq k, \tag{5}$$

и таких, что

$$s_j=\|P_j\|_X^2=\sum_{i=0}^{n} P_j^2(x_i)>0.$$

Будем предполагать, что индекс полинома $P_j(x)$ в точности соответствует его степени.

В силу леммы полиномы $P_j(x)$ $(j=0, 1, 2, \ldots, m)$ линейно независимы на X и, следовательно, линейно независимы также на интервале $(-\infty, \infty)$. Поэтому любой полином $Q_m(x)$ степени не выше m может быть представлен в виде линейной комбинации полиномов (5), т. е.

$$Q_m(x)\equiv b_0P_0(x)+b_1P_1(x)+\ldots+b_mP_m(x), \tag{6}$$

где $b_0, b_1, \ldots, b_m$ — некоторые постоянные числа. Это представление называется *разложением полинома* $Q_m(x)$ *по системе* (4). Коэффициенты $b_0, b_1, \ldots, b_m$ разложения (6) могут быть найдены путем последовательного деления.

Пример 1. Полином $Q_m(x)=x^3$ разложить по системе функций $P_0(x)=1$, $P_1(x)=x-1$, $P_2(x)=x^2-x$, $P_3(x)=x^3-x^2$.

Решение. Деля x^3 на x^3-x^2, имеем

$$x^3\equiv 1\,(x^3-x^2)+x^2.$$

Далее, разделив остаток x^2 на x^2-x, получим

$$x^2\equiv 1\,(x^2-x)+x.$$

Наконец, производя деление нового остатка x на $x-1$, находим $x\equiv 1\,(x-1)+1$.

Таким образом,

$$x^3\equiv 1\,(x^3-x^2)+1\,(x^2-x)+1\,(x-1)+1.$$

В случае, если полиномы $P_j(x)$ $(j=0, 1, \ldots, m)$ ортогональны на множестве точек X, можно дать явные формулы для коэффициен-

тов $b_0, b_1, \ldots, b_m$ разложения (6). Для этого умножим тождество (6) на полином $P_k(x)$ $(k \leqslant m)$ и просуммируем полученное равенство по системе точек $x_0, x_1, \ldots, x_n$. Тогда

$$\sum_{i=0}^{n} Q_m(x_i) P_k(x_i) = b_0 \sum_{i=0}^{n} P_0(x_i) P_k(x_i) + b_1 \sum_{i=0} P_1(x_i) P_k(x_i) + \ldots$$

$$\ldots + b_k \sum_{i=0}^{n} P_k^2(x_i) + \ldots + b_m \sum_{i=0}^{m} P_m(x_i) P_k(x_i).$$

Отсюда, учитывая условия ортогональности (5), находим

$$\sum_{i=0}^{n} Q_m(x_i) P_k(x_i) = b_k \sum_{i=0}^{n} P_k^2(x_i),$$

и, следовательно,

$$b_k = \frac{\sum_{i=0}^{n} Q_m(x_i) P_k(x_i)}{\sum_{i=0}^{n} P_k^2(x_i)} \qquad (k = 0, 1, 2, \ldots, m). \tag{7}$$

Коэффициенты b_k $(k = 0, 1, 2, \ldots, m)$, вычисленные по формулам (7), называются *коэффициентами Фурье полинома* $Q_m(x)$ *относительно данной системы ортогональных на* X *функций* $P_k(x)$ $(k = 0, 1, 2, \ldots, m)$. Если эта система ортонормирована на X, т. е.

$$\sum_{i=0}^{n} P_k^2(x_i) = 1 \qquad (k = 0, 1, 2, \ldots, m),$$

то формулы (7) принимают более простой вид:

$$b_k = \sum_{i=0}^{n} Q_m(x_i) P_k(x_i) \qquad (k = 0, 1, 2, \ldots, m). \tag{7'}$$

Пример 2. Полиномы $P_0(x) = 1$, $P_1(x) = 2x - 5$ и $P_2(x) = 3x^2 - 15x + 10$ ортогональны на системе точек $x_0 = 0$, $x_1 = 1$, $x_2 = 2$, $x_3 = 3$, $x_4 = 4$ и $x_5 = 5$. Разложить полином

$$Q_2(x) = x^3 - 3x + 6$$

по этой системе.

Решение. Полагаем

$$Q_2(x) = b_0 P_0(x) + b_1 P_1(x) + b_2 P_2(x).$$

Запишем результаты вычислений в виде таблицы 5 и таблицы 5а.

Таблица 5

Значения ортогональных полиномов $P_k(x)$ ($k=0, 1, 2$) и их квадратов $P_k^2(x)$ в узлах аппроксимации

x	$P_0(x)$	$P_1(x)$	$P_2(x)$	$P_0^2(x)$	$P_1^2(x)$	$P_2^2(x)$
0	1	—5	10	1	25	100
1	1	—3	—2	1	9	4
2	1	—1	—8	1	1	64
3	1	1	—8	1	1	64
4	1	3	—2	1	9	4
5	1	5	10	1	25	100
Σ				6	70	336

Таблица 5а

Таблица для определения коэффициентов Фурье полинома $Q_2(x)$

x	$Q_2(x)$	$P_0(x)\,Q_2(x)$	$P_1(x)\,Q_2(x)$	$P_2(x)\,Q_2(x)$
0	6	6	—30	60
1	4	4	—12	—8
2	4	4	—4	—32
3	6	6	6	—48
4	10	10	30	—20
5	16	16	80	160
Σ		46	70	112

Отсюда согласно формуле (7) находим: $b_0=46/6=23/3$, $b_1=70/70=1$, $b_2=112/336=1/3$. Следовательно,

$$x^3-3x+6=\frac{23}{3}P_0(x)+P_1(x)+\frac{1}{3}P_2(x).$$

Как легко проверить, тот же результат мы получим, используя метод последовательного деления.

Вернемся к задаче аппроксимирования заданной функции $y=f(x)$ на множестве точек $X=\{x_0, x_1, \ldots, x_n\}$ полиномом данной степени m ($m\leqslant n$).

Искомый полином $Q_m(x)$, для которого квадратичное отклонение

$$S_m=\sum_{i=0}^{n}[Q_m(x_i)-f(x_i)]^2=\min,$$

будем искать в виде обобщенного полинома

$$Q_m(x) = c_0P_0(x) + c_1P_1(x) + \ldots + c_mP_m(x),$$

где полиномы $\{P_k(x)\}$ $(k=0, 1, 2, \ldots, m)$ ортогональны на системе точек X. Отсюда

$$S_m = \sum_{i=0}^{n}[c_0P_0(x_i) + c_1P_1(x_i) + \ldots + c_mP_m(x_i) - f(x_i)]^2.$$

Из необходимого условия минимума S_m для определения коэффициентов c_k $(k=0, 1, 2, \ldots, m)$ имеем систему $m+1$ уравнений

$$\frac{1}{2}\frac{\partial S_m}{\partial c_k} = \sum_{i=0}^{n}[c_0P_0(x_i) + c_1P_1(x_i) + \ldots + c_mP_m(x_i) - f(x_i)]P_k(x_i) = 0 \qquad (k=0, 1, 2, \ldots, m)$$

или

$$c_0\sum_{i=0}^{n}P_0(x_i)P_k(x_i) + c_i\sum_{i=0}^{n}P_1(x_i)P_k(x_i) + \ldots + c_k\sum_{i=0}^{n}P_k^2(x_i) + \ldots + c_m\sum_{i=0}^{n}P_m(x_i)P_k(x_i) = \sum_{i=0}^{n}f(x_i)P_k(x_i)$$
$$(k=0, 1, 2, \ldots, m).$$

В силу ортогональности системы $\{P_j(x)\}$ $(j=0, 1, \ldots, m)$ имеем:

$$\sum_{i=0}^{n}P_j(x_i)P_k(x_i) = 0 \quad \text{при } j \neq k$$

и, кроме того,

$$\sum_{i=0}^{n}P_k^2(x_i) = \|P_k\|_X^2 > 0.$$

Следовательно,

$$c_k = \frac{\sum_{i=0}^{n}f(x_i)P_k(x_i)}{\sum_{i=0}^{n}P_k^2(x_i)} \qquad (k=0, 1, 2, \ldots, m). \tag{8}$$

Таким образом, c_k являются *коэффициентами Фурье функции* $f(x)$ *относительно ортогональной системы* $\{P_k(x)\}$ *на* X.

Для доказательства того, что значения c_j дают минимум величине S_m, составим второй дифференциал. Так как

$$\frac{\partial S_m}{\partial c_k} = 2\left[c_k\sum_{i=0}^{n}P_k^2(x_i) - \sum_{i=0}^{n}f(x_i)P_k(x_i)\right],$$

то

$$\frac{\partial^2 S_m}{\partial c_k \partial c_j} = \begin{cases} 0 \text{ при } k \neq j, \\ 2\sum\limits_{i=0}^{n} P_k^2(x_i) = 2\|P_k\|^2 > 0 \text{ при } k = j. \end{cases}$$

Следовательно,

$$d^2 S_m = \sum_{k=0}^{n} \frac{\partial^2 S_m}{\partial c_k^2} dc_k^2 > 0 \quad \text{при} \quad \sum_{k=0}^{n} dc_k^2 > 0.$$

На основании известной теоремы математического анализа получаем, что при значениях c_k, определяемых формулой (8), величина S_m имеет минимум. Так как этот минимум единственный, то нетрудно видеть, что он является наименьшим значением для S_m при фиксированном m. Полином с коэффициентами Фурье называется *полиномом Фурье* для функции $f(x)$.

Таким образом, можно сформулировать следующий вывод:

Обобщенный полином фиксированного порядка m с коэффициентами Фурье данной функции $f(x)$ на множестве $X = \{x_0, x_1, \ldots, x_n\}$ обладает наименьшим квадратичным отклонением от этой функции на X по сравнению со всеми полиномами того же порядка m.

Вычислим величину S_m для случая, когда c_k являются коэффициентами Фурье. Имеем

$$\begin{aligned} S_m = \sum_{i=0}^{n} [c_0 P_0(x_i) + c_1 P_1(x_i) + \ldots + c_m P_m(x_i) - f(x_i)]^2 = \\ = \sum_{i=0}^{n} [c_0^2 P_0^2(x_i) + c_1^2 P_1^2(x_i) + \ldots + c_m^2 P_m^2(x_i)] + \\ + \sum_{i=0}^{n} f^2(x_i) - 2\sum_{i=0}^{n} [c_0 f(x_i) P_0(x_i) + c_1 f(x_i) P_1(x_i) + \\ + \ldots + c_m f(x_i) P_m(x_i)] + \\ + 2\sum_{i=0}^{n} [c_0 c_1 P_0(x_i) P_1(x_i) + c_0 c_2 P_0(x_i) P_2(x_i) + \\ + \ldots + c_0 c_m P_0(x_i) P_m(x_i) + \ldots + c_m c_{m-1} P_m(x_i) P_{m-1}(x_i)]. \end{aligned}$$

В силу ортогональности системы $[P_k(x)]$ выполнены соотношения

$$\sum_{i=0}^{n} P_k(x_i) P_j(x_i) = 0 \qquad \text{для } k \neq j.$$

Кроме того,

$$c_k = \frac{\sum\limits_{i=0}^{n} f(x_i) P_k(x_i)}{\sum\limits_{i=0}^{n} P_k^2(x_i)}$$

или

$$\sum_{i=0}^{n} f(x_i) P_k(x_i) = c_k \sum_{i=0}^{n} P_k^2(x_i) = c_k \| P_k \|_X^2.$$

Поэтому

$$S_m = c_0^2 \| P_0 \|_X^2 + c_1^2 \| P_1 \|_X^2 + \ldots + c_m^2 \| P_m \|_X^2 + \sum_{i=0}^{n} f^2(x_i) -$$
$$- 2c_0^2 \| P_0 \|_X^2 - 2c_1^2 \| P_1 \|_X^2 - \ldots - 2c_m^2 \| P_m \|_X^2 =$$
$$= \| f(x) \|_X^2 - c_0^2 \| P_0 \|_X^2 - c_1^2 \| P_1 \|_X^2 - \ldots - c_m^2 \| P_m \|_X^2$$

или, в сокращенной записи,

$$S_m = \| f(x) \|_X^2 - \sum_{k=0}^{m} c_k^2 \| P_k \|_X^2.$$

В частности, если система $\{P_k(x)\}$ ортонормирована, то $\| P_k \|_X = 1$ и, следовательно,

$$S_m = \| f(x) \|_X^2 - \sum_{k=0}^{m} c_k^2, \quad \text{где} \quad c_k = \sum_{i=0}^{n} f(x_i) P_k(x_i).$$

§ 6. Полиномы Чебышева, ортогональные на системе равноотстоящих точек

Пусть дана система $n+1$ равноотстоящих точек $X = \{x_0, x_1, \ldots, x_n\}$ с шагом h. С помощью линейного преобразования $t = (x - x_0)/h$ переведем эти точки соответственно в $t = 0, 1, 2, \ldots, n$.

Полиномы $P_{0,n}(t)$, $P_{1,n}(t)$, $\ldots$, $P_{m,n}(t)$ $(m \leqslant n)$ соответственно степеней $0, 1, \ldots, m$, ортогональные на множестве точек $\{0, 1, 2, \ldots, n\}$ и отличные от нуля на этом множестве, называются *ортогональными полиномами Чебышева*.

Заметим, что в обозначении полинома Чебышева $P_{k,n}(t)$ первый индекс k представляет собой степень полинома, а второй индекс n — число точек, уменьшенное на единицу.

Полиномы Чебышева могут быть заданы формулой [3]:

$$P_{k,n}(t) = \sum_{s=0}^{n} (-1)^s C_k^s C_{k+s}^s \frac{t^{[s]}}{n^{[s]}} \qquad (1)$$
$$(k = 0, 1, 2, \ldots, m),$$

где $t^{[s]} = t(t-1)\ldots(t-s+1)$ и $n^{[s]} = n(n-1)\ldots(n-s+1)$ соответствующие обобщенные степени.

Запишем несколько первых ортогональных полиномов Чебышева:

$$\left.\begin{aligned} P_{0,n}(t)=1,\ P_{1,n}(t)=1-2\frac{t}{n} \quad (n\geqslant 1),\\ P_{2,n}(t)=1-6\frac{t}{n}+6\frac{t(t-1)}{n(n-1)} \quad (n\geqslant 2),\\ P_{3,n}(t)=1-12\frac{t}{n}+30\frac{t(t-1)}{n(n-1)}-20\frac{t(t-1)(t-2)}{n(n-1)(n-2)} \quad (n\geqslant 3).\end{aligned}\right\} \tag{2}$$

Возвращаясь к прежней переменной x, получим систему полиномов

$$P_{k,n}\left(\frac{x-x_0}{n}\right) \quad (k=0,\ 1,\ \ldots,\ m;\ m\leqslant n),$$

ортогональных на множестве X.

Система полиномов $\{P_{k,n}(t)\}$ не является нормированной. Можно показать [3], что

$$\|P_{k,n}(t)\|^2=\sum_{i=0}^{n}P_{k,n}^2(i)=\frac{(n+k+1)^{[k+1]}}{(2k+1)\,n^{[k]}}. \tag{3}$$

Разделив многочлены $P_{k,n}(t)$ на их нормы, мы получим *нормированную систему ортогональных полиномов Чебышева*

$$\tilde{P}_{k,n}(t)=\frac{P_{k,n}(t)}{\|P_{k,n}(t)\|} \quad (k=0,\ 1,\ 2,\ \ldots,\ m;\ m\leqslant n). \tag{4}$$

Пример 1. Получить систему полиномов до третьей степени включительно, ортонормированных на системе точек $x_0=1/2$; $x_1=1$; $x_2=3/2$; $x_3=2$; $x_4=5/2$; $x_5=2$.

Решение. Полагая

$$t=\frac{x-x_0}{n}=\frac{x-\frac{1}{2}}{\frac{1}{2}}=2\left(x-\frac{1}{2}\right),$$

переведем точки x_i в целочисленные точки $t=0,\ 1,\ 2,\ 3,\ 4,\ 5$. Теперь в формулах (1) принимаем $n=5$.

Имеем:

$$P_{0,5}(t)=1,\qquad P_{1,5}(t)=1-0{,}4t,\qquad P_{2,5}(t)=1-1{,}2t+0{,}3t(t-1),$$
$$P_{3,5}(t)=1-2{,}4t+1{,}5t(t-1)-0{,}333t(t-1)(t-2).$$

Нормы функций $P_{k,5}(t)$ $(k=0,\ 1,\ 2,\ 3)$ вычисляем по формуле (3):

$$\|P_{0,5}(t)\|=\sqrt{6},\qquad \|P_{1,5}(t)\|=\sqrt{\frac{7\cdot 6}{3\cdot 5}}=\sqrt{\frac{14}{5}},$$
$$\|P_{2,5}(t)\|=\sqrt{\frac{8\cdot 7\cdot 6}{5\cdot 5\cdot 4}}=\frac{2}{5}\sqrt{21},$$
$$\|P_{3,5}(t)\|=\sqrt{\frac{9\cdot 8\cdot 7\cdot 6}{7\cdot 5\cdot 4\cdot 3}}=\frac{6}{\sqrt{5}}.$$

Разделив полиномы $P_{k,5}(t)$ на их нормы и переходя от переменной t к переменной x, получим нормированную систему ортогональных полиномов Чебышева:

$$\tilde{P}_{0,5}(x)=\frac{1}{\sqrt{6}},\quad \tilde{P}_{1,5}(x)=\sqrt{\frac{5}{14}}\left[1-0{,}8\left(x-\frac{1}{2}\right)\right],$$

$$\tilde{P}_{2,5}(x)=\frac{5}{2\sqrt{21}}\left[1-2{,}4\left(x-\frac{1}{2}\right)+0{,}6\left(x-\frac{1}{2}\right)\left(x-\frac{3}{2}\right)\right],$$

$$\tilde{P}_{3,5}(x)=\frac{\sqrt{5}}{6}\left[1-4{,}8\left(x-\frac{1}{2}\right)+3\left(x-\frac{1}{2}\right)\left(x-\frac{3}{2}\right)-\right.$$

$$\left.-0{,}666\left(x-\frac{1}{2}\right)\left(x-\frac{3}{2}\right)\left(x-\frac{5}{2}\right)\right].$$

Если функция $y=f(x)$ задана на множестве узлов $X=\{x_0, x_1, \ldots, x_n\}$ с шагом h, то наилучший квадратично аппроксимирующий ее на X полином степени m будет иметь вид

$$Q_m(x)=\sum_{k=0}^{m} c_k P_{k,n}\left(\frac{x-x_0}{h}\right), \tag{5}$$

где

$$c_k=\frac{\sum_{i=0}^{n} f(x_i)\,P_{k,n}(i)}{\|P_{k,n}(t)\|^2}\qquad (k=0,\ 1,\ 2,\ \ldots,\ m) \tag{6}$$

— коэффициенты Фурье функции $f(x)$ относительно системы ортогональных полиномов Чебышева $P_{k,n}((x-x_0)/h)$ $(k=0,\ 1,\ 2,\ \ldots,\ m)$.

Из формул (5) и (6) следует, что полином $Q_m(x)$ не изменится, если ортогональные полиномы Чебышева $P_{k,n}((x-x_0)/h)$ $(k=0, 1, 2, \ldots, m)$ умножить на некоторые постоянные множители, отличные от нуля. Этим обстоятельством пользуются при вычислениях, чтобы получать целые значения модифицированных ортогональных полиномов Чебышева в узлах аппроксимирования.

В частности, если воспользоваться нормированными ортогональными полиномами Чебышева $\tilde{P}_{k,n}(t)$, то будем иметь

$$Q_m(x)=\sum_{k=0}^{m} \tilde{c}_k \tilde{P}_{k,n}\left(\frac{x-x_0}{h}\right), \tag{7}$$

где

$$\tilde{c}_k=\sum_{i=0}^{n} f(x_i)\,\tilde{P}_{k,n}(i)\quad (k=0,\ 1,\ \ldots,\ m). \tag{8}$$

Пример 2. Получить по способу наименьших квадратов полином пятой степени, приближенно представляющий табличную функцию [3]:

x	0	0,3	0,6	0,9	1,2	1,5	1,8	2,1	2,4	2,7
y	1,300	1,245	1,095	0,855	0,514	0,037	—0,600	—1,295	—1,767	—1,914

Решение. Здесь $m=5$, $n=9$. Полагаем $t=x/0{,}3$.

Для удобства выкладок воспользуемся полиномами

$$\hat{P}_{k,n}(t)=\rho_{k,n}\,P_{k,n}(t),$$

отличающимися от ортогональных полиномов Чебышева $P_{k,n}(t)$ лишь числовыми множителями $\rho_{k,n}=\hat{P}_{k,n}(0)$, подобранными так, чтобы для целочисленных значений аргумента t значения $\hat{P}_{k,n}(t)$ были также целочисленными.

Полагаем

$$Q(t)=\sum_{k=0}^{5}\hat{b}_k\hat{P}_{k,9}(t),$$

где $\hat{b}_k=\frac{\hat{c}_k}{\hat{s}_k}$, $\hat{c}_k=\sum_{i=0}^{9}y_i\hat{P}_{k,9}(i)$, $\hat{s}_k=\sum_{i=0}^{9}\hat{P}^2_{k,9}(i)$. Значения полиномов $\hat{P}_{k,9}(t)$ приведены в приложении 1.

Вычисления расположим так, как показано в таблице 6.

Отсюда аппроксимирующий полином имеет вид

$$Q(t)=-0{,}053+0{,}202\hat{P}_{1,9}(t)-0{,}057\hat{P}_{2,9}(t)-0{,}005\hat{P}_{3,9}(t)+$$
$$+0{,}005\hat{P}_{4,9}(t)-0{,}002\hat{P}_{5,9}(t)$$

или

$$Q(t)=-0{,}053+1{,}818P_{1,9}(t)-0{,}342P_{2,9}(t)-0{,}210P_{3,9}(t)+$$
$$+0{,}090P_{4,9}(t)-0{,}012P_{5,9}(t).$$

Пример 3. По способу наименьших квадратов получить полином третьей степени, аппроксимирующий функцию $y=f(x)$, заданную таблицей

x	0	0,05	0,10	0,15	0,20	0,25	0,30	0,35	0,40	0,45	0,50
y	0	0,040	0,080	0,119	0,158	0,197	0,236	0,274	0,311	0,347	0,383

Решение. Искомый полином ищем в виде

$$Q_3(x)\equiv\tilde{Q}_3(t)=\tilde{c}_0\tilde{P}_{0,10}(t)+\tilde{c}_1\tilde{P}_{1,10}(t)+\tilde{c}_2\tilde{P}_{2,10}(t)+\tilde{c}_3\tilde{P}_{3,10}(t), \qquad (9)$$

где $t=(x-x_0)/h=x/0{,}05=20x$ и $\tilde{P}_{k,10}(t)$ $(k=0,\ 1,\ 2,\ 3)$ — полиномы Чебышева (4), ортонормированные на множестве точек $\{0,\ 1,\ 2,\ \ldots,\ 10\}$.

Коэффициенты $\tilde{c}_0$, $\tilde{c}_1$, $\tilde{c}_2$ и $\tilde{c}_3$ в равенстве (9) находим по формуле

$$\tilde{c}_m=\sum_{i=0}^{10}y_i\tilde{P}_{m,10}(i) \qquad (m=0,\ 1,\ 2,\ 3), \qquad (10)$$

где $y_i=f(x_i)$.

Таблица 6

Построение аппроксимирующего полинома методом Чебышева

t	x	y	$P_0(t)$	$\hat{P}_{1,9}(t)$	$\hat{P}_{2,9}(t)$	$\hat{P}_{3,9}(t)$	$\hat{P}_{4,9}(t)$	$\hat{P}_{5,9}(t)$	$\hat{y}$	$\varepsilon=y-\hat{y}$
0	0	1,300	1	9	6	42	18	6	1,310	−0,010
1	0,3	1,245	1	7	2	−14	−22	−14	1,236	+0,009
2	0,6	1,095	1	5	−1	−35	−17	1	1,098	−0,003
3	0,9	0,855	1	3	−3	−31	3	11	0,868	−0,013
4	1,2	0,514	1	1	−4	−12	18	6	0,514	0
5	1,5	0,037	1	−1	−4	12	18	−6	0,017	+0,020
6	1,8	−0,600	1	−3	−3	34	3	−11	−0,002	+0,002
7	2,1	−1,295	1	−5	−1	35	−17	−1	−1,263	−0,1032
8	2,4	−1,767	1	−7	2	14	−22	14	−1,793	+0,026
9	2,7	−1,914	1	−9	6	−42	18	−6	−1,908	−0,006
		$\hat{c}_i$	−0,530	66,802	−7,497	−14,659	14,515	−1,627		
		$\hat{s}_i$	10	330	132	8850	2560	780		
		$\hat{b}_i$	−0,053	0,202	−0,057	−0,005	0,005	−0,002		

Значения ортонормированных полиномов Чебышева в целочисленных точках ($t=0, 1, 2, \ldots, 10$) берем из соответствующих таблиц *). Выпишем эти значения (таблица 7) для $n=10$ и $m=1, 2, 3$, оставляя в них 4 знака после запятой (таблицы десятичные). Построение аппроксимирующего полинома приведено в таблице 7а.

Таблица 7

Значения ортонормированных полиномов Чебышева

t	$\tilde{P}_{1,10}(t)$	$\tilde{P}_{2,10}(t)$	$\tilde{P}_{3,10}(t)$
0	0,4767	0,5121	0,4580
1	0,3814	0,2048	—0,0916
2	0,2860	—0,0341	—0,3359
3	0,1907	—0,2048	—0,3516
4	0,0953	—0,3073	—0,2137
5	0,0000	—0,3414	0,0000
6	—0,0953	—0,3073	0,2137
7	—0,1907	—0,2048	0,3516
8	—0,2860	—0,0341	0,3359
9	—0,3814	0,2048	0,0916
10	—0,4767	0,5121	—0,4580

Таблица 7а

Построение аппроксимирующего полинома с помощью ортонормированных полиномов Чебышева

t	$y=f(t)$	$y\tilde{P}_{1,10}(t)$	$y\tilde{P}_{2,10}(t)$	$y\tilde{P}_{3,10}(t)$	$\tilde{y}$	$\varepsilon=y-\tilde{y}$
0	0	0	0	0	0	0
1	0,040	0,0153	0,0082	—0,0036	0,0399	0,0001
2	0,080	0,0229	—0,0027	—0,0269	0,0796	0,0004
3	0,119	0,0227	—0,0244	—0,0418	0,1191	0,0009
4	0,158	0,0151	—0,0486	—0,0338	0,1592	—0,0012
5	0,197	0	—0,0673	0	0,1970	0,000
6	0,236	—0,0225	—0,0725	0,0504	0,2352	0,0008
7	0,274	—0,0523	—0,0561	0,0963	0,2727	0,0013
8	0,311	—0,0896	—0,0106	0,1045	0,3097	0,0013
9	0,347	—0,1323	—0,0711	0,0318	0,3406	0,0094
10	0,383	—0,1826	0,1961	—0,1754	0,3708	0,0092
$\sum$	2,145	—0,4033	—0,0066	0,0014		

*) Л. П. Грабарь, Таблицы полиномов Чебышева, ортонормированных на системе равноотстоящих точек. Вычислительный центр АН СССР, Москва, 1965.

Отсюда, учитывая, что $\tilde{P}_{0,10}(t)=1$ и $\tilde{P}_{0,10}(t)=1/\|P_{0,10}(t)\|=1/\sqrt{11}$, будем иметь

$$\tilde{c}_0\tilde{P}_{0,10}(t)=\sum_{i=0}^{n} y_i \frac{1}{\sqrt{11}}\frac{1}{\sqrt{11}}=\frac{1}{11}\sum_{i=0}^{n} y_i=0{,}0909\sum_{i=0}^{10} y_i,$$

и, следовательно,

$$\tilde{Q}_3(t)=0{,}1950-0{,}4033\tilde{P}_{0,10}(t)-0{,}0066\tilde{P}_{2,10}(t)+0{,}0014\tilde{P}_{3,10}(t). \qquad (11)$$

Так как

$$\tilde{P}_{k,10}(t)=\frac{P_{k,10}(t)}{\|P_{k,10}\|},$$

где

$$\|P_{1,10}\|=2{,}0976, \quad \|P_{2,10}\|=1{,}9528, \quad \|P_{3,10}\|=2{,}1833,$$

то имеем

$$\tilde{Q}_3(t)=0{,}1950-0{,}1919P_{1,10}(t)-0{,}0034P_{2,10}(t)+0{,}0006P_{3,10}(t)=$$

$$=0{,}1950-0{,}1919\left(1-\frac{t}{5}\right)-0{,}0034\left[1-\frac{3t}{5}+\frac{t(t-1)}{15}\right]+$$

$$+0{,}0006\left[1-\frac{6t}{5}+\frac{t(t-1)}{3}-\frac{t(t-1)(t-2)}{36}\right]. \qquad (12)$$

Полагая $t=20x$, окончательно получаем

$$Q_3(x)=0{,}1333x^3+0{,}0093x^2+0{,}7938x+0{,}0003.$$

В последних двух столбцах таблицы 7а приведены значения полинома $\tilde{y}=\tilde{Q}_3(t)$ в точках $t=0, 1, \ldots, 10$, вычисленные по формуле (12), и соответствующие разности $\varepsilon=y-\tilde{y}$.

§ 7. Интегральное квадратичное аппроксимирование функций на отрезке

В интегральном способе наименьших квадратов за меру отклонения функции $f(x)$ от полинома

$$c_0+c_1x+\ldots+c_mx^m=Q_m(x)$$

на отрезке $[a, b]$ принимается величина

$$I_m=\int_a^b [Q_m(x)-f(x)]^2\,dx.$$

Очевидно, I_m есть функция коэффициентов $c_0, c_1, \ldots, c_m$. Для наилучшей квадратичной аппроксимации эти коэффициенты нужно подобрать

так, чтобы I_m принимало свое наименьшее значение. Отсюда получаем систему уравнений

$$\left.\begin{aligned}
\frac{1}{2}\frac{\partial I_m}{\partial c_0} &= \int_a^b [Q_m(x)-f(x)]\,dx = 0,\\
\frac{1}{2}\frac{\partial I_m}{\partial c_1} &= \int_a^b [Q_m(x)-f(x)]\,x\,dx = 0,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\\
\frac{1}{2}\frac{\partial I_m}{\partial c_m} &= \int_a^b [Q_m(x)-f(x)]\,x^m\,dx = 0
\end{aligned}\right\} \qquad (1)$$

или, учитывая, что

$$s_k = \int_a^b x^k\,dx = \frac{b^{k+1}-a^{k+1}}{k+1} \qquad (k=0,\ 1,\ 2,\ \ldots,\ m),$$

будем иметь:

$$\left.\begin{aligned}
s_0c_0 + s_1c_1 + \ldots + s_mc_m &= \int_a^b f(x)\,dx,\\
s_1c_0 + s_2c_1 + \ldots + s_{m+1}c_m &= \int_a^b xf(x)\,dx,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\\
s_mc_0 + s_{m+1}c_1 + \ldots + s_{2m}c_m &= \int_a^b x^m f(x)\,dx.
\end{aligned}\right\} \qquad (2)$$

Нетрудно убедиться, что система (2) имеет единственное решение. Доказывается [2], что решения $c_0, c_1, \ldots, c_m$ дают наименьшее значение величине I_m.

Пример 1. Найти наилучшую квадратичную аппроксимацию посредством полинома второй степени функции $f(x)=\sqrt{x}$ на отрезке [0, 1].

Решение. Так как

$$\int_0^1 x^k\,dx = \frac{1}{k+1},$$

то коэффициенты полинома $Q(x)=c_0+c_1x+c_2x^2$ определяются из системы

$$\left.\begin{aligned}
c_0 + \frac{1}{2}c_1 + \frac{1}{3}c_2 &= \frac{2}{3},\\
\frac{1}{2}c_0 + \frac{1}{3}c_1 + \frac{1}{4}c_2 &= \frac{2}{5},\\
\frac{1}{3}c_0 + \frac{1}{4}c_1 + \frac{1}{5}c_2 &= \frac{2}{7}.
\end{aligned}\right\}$$

Отсюда $c_0=6/35$, $c_1=48/35$, $c_2=-4/7$, и, следовательно (рис. 3),

$$Q(x)=\frac{6}{35}+\frac{48}{35}x-\frac{4}{7}x^2.$$

Пример 2. Функцию 3^x квадратично аппроксимировать на отрезке $[-1, 1]$ полиномом третьей степени.

Решение. Полагая $Q_3(x)=c_0+c_1x+c_2x^2+c_3x^3$, для коэффициентов полинома $Q_3(x)$ получаем систему уравнений

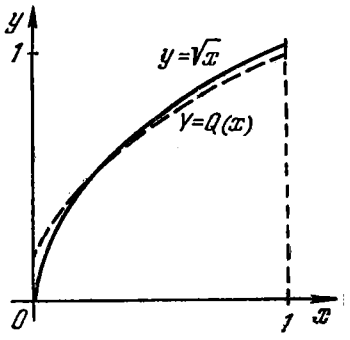

Рис. 3.

$$\left.\begin{aligned}
2c_0+\frac{2}{3}c_2&=\int_{-1}^{1}3^x\,dx,\\
\frac{2}{3}c_1+\frac{2}{5}c_3&=\int_{-1}^{1}x\cdot 3^x\,dx,\\
\frac{2}{3}c_0+\frac{2}{5}c_2&=\int_{-1}^{1}x^2\cdot 3^x\,dx,\\
\frac{2}{5}c_1+\frac{2}{7}c_3&=\int_{-1}^{1}x^3\cdot 3^x\,dx
\end{aligned}\right\}$$

или

$$6c_0+2c_2=7{,}2819,\quad 2c_1+1{,}2c_2=2{,}4739,\quad 2c_0+1{,}2c_2=2{,}7779,\quad 2{,}8c_1+2c_3=3{,}5366.$$

Отсюда находим: $c_0=0{,}9944$, $c_1=1{,}1000$, $c_2=0{,}6576$ и $c_3=0{,}2335$.

Следовательно,

$$Q(x)=0{,}9944+1{,}1000x+0{,}6576x^2+0{,}2335x^3.$$

В общем случае данную непрерывную функцию $f(x)$ можно аппроксимировать на конечном отрезке $[a, b]$ с помощью обобщенных полиномов

$$Q_m(x)=c_0\varphi_0(x)+c_1\varphi_1(x)+\ldots+c_m\varphi_m(x),$$

где $\{\varphi_i(x)\}$ — заданная система непрерывных функций и c_i — постоянные коэффициенты.

Согласно способу наименьших квадратов коэффициенты c_i $(i=0, 1, 2, \ldots, m)$ подбираются так, чтобы квадратичное отклонение полинома $Q(x)$ от функции $f(x)$, равное

$$I_m=\int_a^b[Q(x)-f(x)]^2\,dx=\int_a^b\Big[\sum_{i=0}^{m}c_i\varphi_i(x)-f(x)\Big]^2dx,$$

имело наименьшее значение.

Для нахождения минимума функции $I_m=I_m(c_0, c_1, \ldots, c_m)$, как известно, нужно составить все частные производные

$$\frac{\partial I_m}{\partial c_i}\qquad (i=0, 1, 2, \ldots, m)$$

и приравнять их нулю. Это дает систему для определения коэффициентов $c_0, c_1, \ldots, c_n$:

$$\left.\begin{aligned}
&\frac{1}{2}\frac{\partial I_m}{\partial c_0}=\int_a^b\left[\sum_{i=0}^m c_i\varphi_i(x)-f(x)\right]\varphi_0(x)\,dx=0,\\
&\frac{1}{2}\frac{\partial I_m}{\partial c_1}=\int_a^b\left[\sum_{i=0}^m c_i\varphi_i(x)-f(x)\right]\varphi_1(x)\,dx=0,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&\frac{1}{2}\frac{\partial I_m}{\partial c_m}=\int_a^b\left[\sum_{i=0}^m c_i\varphi_i(x)-f(x)\right]\varphi_m(x)\,dx=0.
\end{aligned}\right\}\qquad(3)$$

Вводя сокращенные обозначения

$$(\varphi,\ \psi)=\int_a^b\varphi(x)\,\psi(x)\,dx,$$

систему (3) можно записать в следующем виде:

$$\left.\begin{aligned}
&c_0(\varphi_0,\ \varphi_0)+c_1(\varphi_1,\ \varphi_0)+\ldots+c_n(\varphi_n,\ \varphi_0)=(f,\ \varphi_0),\\
&c_0(\varphi_0,\ \varphi_1)+c_1(\varphi_1,\ \varphi_1)+\ldots+c_n(\varphi_n,\ \varphi_1)=(f,\ \varphi_1),\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&c_0(\varphi_0,\ \varphi_n)+c_1(\varphi_1,\ \varphi_n)+\ldots+c_n(\varphi_n,\ \varphi_n)=(f,\ \varphi_n).
\end{aligned}\right\}\qquad(4)$$

Доказывается [2], что если функции $\varphi_0(x), \varphi_1(x), \ldots, \varphi_n(x)$ линейно независимы на $[a, b]$, то система (4) имеет единственное решение, которое соответствует наименьшему квадратичному отклонению I_m.

Заметим, что неудобством интегральной квадратичной аппроксимации является необходимость вычисления определенных интегралов, которые могут быть весьма сложными и даже не выражаться через элементарные функции. В этом смысле способ точечной квадратичной аппроксимации предпочтительнее.

§ 8. Ортогональные на промежутке системы функций

В случае, когда степень аппроксимирующего полинома велика, метод отыскания такого полинома, указанный в § 6, становится громоздким и его заменяют другим методом, основанным на идее ортогональных функций.

Определение. Система интегрируемых функций

$$\varphi_0(x),\ \varphi_1(x),\ \ldots,\ \varphi_n(x)\qquad(1)$$

называется *ортогональной* на $[a, b]$, если

$$(\varphi_m, \varphi_n) = \int_a^b \varphi_m(x)\,\varphi_n(x)\,dx = 0 \text{ при } m \neq n.$$

Число

$$\|\varphi_m(x)\| = \sqrt{(\varphi_m, \varphi_m)} = \sqrt{\int_a^b \varphi_m^2(x)\,dx}$$

называется *нормой* функции $\varphi_m(x)$ на отрезке $[a, b]$.

Если нормы всех функций системы (1) равны единице, то эта система называется *ортонормированной* (ортогональной и нормированной). Для ортонормированной системы выполняется условие

$$\int_a^b \varphi_m(x)\,\varphi_n(x)\,dx = \delta_{mn},$$

где δ_{mn} — *символ Кронекера*, т. е. $\delta_{mn} = 0$ при $m \neq n$; $\delta_{mn} = 1$ при $m = n$.

Очевидно, всякую систему, не содержащую функций с нулевой нормой, можно нормировать. Для этого достаточно каждую функцию разделить на ее норму. Система функций

$$\psi_0(x) = \frac{\varphi_0(x)}{\|\varphi_0(x)\|};\quad \psi_1(x) = \frac{\varphi_1(x)}{\|\varphi_1(x)\|};\quad \ldots;\quad \psi_n(x) = \frac{\varphi_n(x)}{\|\varphi_n(x)\|};\ldots$$

нормирована, так как

$$\int_a^b \psi_i^2(x)\,dx = \int_a^b \frac{\varphi_i^2(x)}{\|\varphi_i(x)\|^2}\,dx = \frac{1}{\|\varphi_i(x)\|^2}\int_a^b \varphi_i^2(x)\,dx = 1.$$

Пример 1. Пронормировать систему функций

$$1,\ x,\ x^2,\ \ldots,\ x^m,\ \ldots, \tag{2}$$

заданную на отрезке $[0,1]$.

Решение. Имеем

$$\psi_0(x) = \frac{1}{\sqrt{\int_0^1 1^2\,dx}} = 1,$$

$$\psi_m(x) = \frac{x^m}{\sqrt{\int_0^1 x^{2m}\,dx}} = \frac{x^m}{\sqrt{\frac{1}{2m+1}}} = \sqrt{2m+1}\,x^m \quad (m = 1, 2, \ldots).$$

После нормирования система (2) имеет вид

$$1,\ \sqrt{3}x,\ \sqrt{5}x^2,\ \ldots,\ \sqrt{2m+1}\,x^m,\ \ldots$$

Если система функций $\{\varphi_i(x)\}$ ортогональна на отрезке $[a, b]$, то задача о квадратичной аппроксимации данной непрерывной функции $f(x)$ на заданном отрезке $[a, b]$ с помощью обобщенного полинома

$$Q_m(x) = c_0\varphi_0(x) + c_1\varphi_1(x) + \ldots + c_m\varphi_m(x)$$

получает простое решение.

В самом деле, из необходимого условия минимума интеграла

$$I_m = \int_a^b \left[\sum_{i=0}^{m} c_i\varphi_i(x) - f(x)\right]^2 dx \tag{3}$$

для определения коэффициентов c_i $(i = 0, 1, 2, \ldots, m)$ имеем систему $m+1$ уравнений

$$\frac{1}{2}\frac{\partial I_m}{\partial c_j} = \int_a^b \left[\sum_{i=0}^{m} c_i\varphi_i(x) - f(x)\right] \varphi_j(x)\,dx = 0 \quad (j = 0, 1, 2, \ldots, m). \tag{4}$$

После несложных упрощений система (4) принимает вид

$$\sum_{i=0}^{m} c_i \int_a^b \varphi_i(x)\,\varphi_j(x)\,dx = \int_a^b f(x)\,\varphi_j(x)\,dx. \tag{5}$$

В силу ортогональности системы $\{\varphi_i(x)\}$ выполнено равенство

$$\int_a^b \varphi_i(x)\,\varphi_j(x)\,dx = 0 \quad \text{при} \quad i \neq j.$$

Поэтому все слагаемые левой части уравнений (5), за исключением j-го, обращаются в нуль. Следовательно,

$$c_j \int_a^b \varphi_j^2(x)\,dx = \int_a^b f(x)\,\varphi_j(x)\,dx.$$

Предполагая, что среди функций $\varphi_i(x)$ $(i = 0, 1, 2, \ldots, m)$ нет функции с нулевой нормой, т. е.

$$\|\varphi_j(x)\|^2 = \int_a^b \varphi_j^2(x)\,dx > 0,$$

получим окончательно

$$c_j = \frac{\int_a^b f(x)\,\varphi_j(x)\,dx}{\int_a^b \varphi_j^2(x)\,dx} \quad (j = 0, 1, 2, \ldots, m). \tag{6}$$

В знаменателе формулы (6) стоит квадрат нормы функции $\varphi_j(x)$, т. е.

$$c_j = \frac{\int_a^b f(x)\,\varphi_j(x)\,dx}{\|\varphi_j(x)\|^2}. \tag{7}$$

В случае ортонормированной системы коэффициенты вычисляются особенно просто:

$$c_j = \int_a^b f(x)\,\varphi_j(x)\,dx$$
$$(j = 0,\ 1,\ 2,\ \ldots,\ m). \tag{8}$$

Коэффициенты c_j, определяемые формулой (7), называются *коэффициентами Фурье функции* $f(x)$ относительно заданной ортогональной системы $\{\varphi_i(x)\}$.

Для доказательства того, что значения c_j дают минимум интегралу I_m, составим второй дифференциал d^2I_m. Из формулы (4) имеем

$$\frac{1}{2}\frac{\partial^2 I_m}{\partial c_j \partial c_k} = \int_a^b \varphi_j(x)\,\varphi_k(x)\,dx = 0 \text{ при } j \neq k.$$

Следовательно,

$$d^2I_m = \sum_{j=0}^{m} \frac{\partial^2 I_m}{\partial c_j^2}\,dc_j^2 + 2\sum_{j<k} \frac{\partial^2 I_m}{\partial c_j \partial c_k}\,dc_j dc_k = 2\sum_{j=0}^{m} \|\varphi_j(x)\|^2\,dc_j^2 > 0$$

при

$$\sum_{j=0}^{m} dc_j^2 \neq 0.$$

На основании известной теоремы математического анализа получаем, что при значениях c_j, определяемых формулой (7), интеграл I_m имеет минимум. Более того, так как интеграл I_m имеет единственный минимум, то значение I_m, соответствующее коэффициентам Фурье c_j, является наименьшим в пространстве коэффициентов $c_0,\ c_1,\ \ldots,\ c_m$.

Таким образом, можно сформулировать следующий вывод: *обобщенный полином с коэффициентами Фурье данной функции обладает наименьшим квадратичным отклонением от этой функции по сравнению со всеми другими обобщенными полиномами того же порядка* m (ср. § 5).

Дадим оценку отклонения (3) для случая, когда c_i являются коэффициентами Фурье. Имеем

$$I_m = \int_a^b [f(x) - \sum_{i=0}^m c_i\varphi_i(x)]^2\,dx =$$
$$= \int_a^b [f^2(x) - 2f(x) \sum_{i=0}^m c_i\varphi_i(x) + \sum_{i=0}^m c_i^2\varphi_i^2(x) +$$
$$+ 2\sum_{i<k}^m c_ic_k\varphi_i(x)\,\varphi_k(x)]\,dx.$$

Переходя от интеграла суммы к сумме интегралов, получим

$$I_m = \int_a^b f^2(x)\,dx - 2\sum_{i=0}^m c_i \int_a^b f(x)\,\varphi_i(x)\,dx + \sum_{i=0}^m c_i^2 \int_a^b \varphi_i^2(x)\,dx +$$
$$+ 2\sum_{i<k}^m c_ic_k \int_a^b \varphi_i(x)\,\varphi_k(x)\,dx. \qquad (9)$$

Так как

$$\int_a^b f(x)\,\varphi_i(x)\,dx = c_i \int_a^b \varphi_i^2(x)\,dx \quad (i = 0,\ 1,\ 2,\ \ldots,\ m)$$

и

$$\int_a^b \varphi_k(x)\,\varphi_i(x)\,dx = 0 \quad \text{при} \quad i \neq k,$$

то после приведения подобных членов в формуле (9) имеем

$$I_m = \int_a^b f^2(x)\,dx - \sum_{i=0}^m c_i^2 \int_a^b \varphi_i^2(x)\,dx \qquad (10)$$

или, в терминах нормы функций,

$$I_m = \|f(x)\|^2 - \sum_{i=0}^m c_i^2\,\|\varphi_i(x)\|^2. \qquad (11)$$

Так как $I_m \geqslant 0$, то из формулы (11) получается так называемое *неравенство Бесселя*

$$\sum_{i=0}^m c_i^2\,\|\varphi_i(x)\|^2 \leqslant \|f(x)\|^2. \qquad (12)$$

В частности, при $m \longrightarrow \infty$ получаем

$$\sum_{i=0}^\infty c_i^2\,\|\varphi_i(x)\|^2 \leqslant \|f(x)\|^2.$$

Если система $\varphi_0(x)$, $\varphi_1(x)$, ..., $\varphi_m(x)$, ... ортонормирована, то формулы (11) и (12) упрощаются. В этом случае

$$I_m = \| f(x) \|^2 - \sum_{i=0}^{m} c_i^2 \tag{13}$$

и

$$\sum_{i=0}^{m} c_i^2 \leqslant \| f(x) \|^2. \tag{14}$$

Отметим, что обобщенный полином $Q_m(x)$ с коэффициентами Фурье обладает важными свойствами:

1) при увеличении числа слагаемых m младшие коэффициенты c_j, как следует из формулы (7), остаются неизменными, т. е. при добавлении новых членов проделанная прежде работа сохраняется полностью;

2) при увеличении m квадратичная погрешность

$$I_m = \int_a^b [Q_m(x) - f(x)]^2 \, dx$$

в силу формулы (10) монотонно убывает в широком смысле, т. е. $I_1 \geqslant I_2 \geqslant \ldots \geqslant I_m \geqslant I_{m+1} \geqslant \ldots$ Таким образом, присоединение новых слагаемых увеличивает точность аппроксимации.

Аналогичные свойства для полиномов Фурье имеют место также в случае точечной аппроксимации.

Если система ортогональных функций $\{\varphi_i(x)\}$ такова, что для *любой непрерывной* функции $f(x)$ справедливо соотношение

$$\lim_{m \to \infty} I_m = 0,$$

то эта система называется *полной*. В противном случае система называется *неполной*.

Предполагая, что система $\{\varphi_i(x)\}$ полная, и переходя к пределу в равенстве (11), будем иметь

$$0 = \| f(x) \|^2 - \sum_{i=0}^{\infty} c_i^2 \| \varphi_i(x) \|^2, \text{ т. е. } \sum_{i=0}^{\infty} c_i^2 \| \varphi_i(x) \|^2 = \| f(x) \|^2. \tag{15}$$

В частности, если полная система ортонормирована, то

$$\sum_{i=0}^{\infty} c_i^2 = \| f(x) \|^2. \tag{16}$$

Равенство (15) называется *равенством Парсеваля* или *условием полноты*. Если c_i рассматривать как компоненты функции $f(x)$ относительно ортонормированной системы $\{\varphi_i(x)\}$, то можно сказать, что равенство Парсеваля (16) является аналогом теоремы Пифагора в функциональном пространстве.

§ 9. Понятие о гармоническом анализе

В качестве примера ортогональной системы рассмотрим тригонометрическую систему

$$1,\ \sin x,\ \cos x,\ \sin 2x,\ \cos 2x,\ \ldots,\ \sin nx,\ \cos nx,\ \ldots \qquad (1)$$

Покажем, что эта система ортогональна на любом отрезке длины 2π, например на отрезке $[-\pi, \pi]$. Для этого проверим обращение в нуль интегралов:

I. $\int\limits_{-\pi}^{\pi} \sin mx \sin nx\, dx$, если m и n — целые и $m \neq n$.

II. $\int\limits_{-\pi}^{\pi} \cos mx \cos nx\, dx$, если m и n — целые и $m \neq n$.

III. $\int\limits_{-\pi}^{\pi} \cos mx \sin nx\, dx$ для всех целых m и n.

Действительно,

$$\text{I.}\quad \int\limits_{-\pi}^{\pi} \sin mx \sin nx\, dx = \frac{1}{2}\left[\frac{\sin(m-n)x}{m-n} - \frac{\sin(m+n)x}{m+n}\right]_{-\pi}^{\pi} = 0,$$

если m и n — целые и $m \neq n$.

$$\text{II.}\quad \int\limits_{-\pi}^{\pi} \cos mx \cos nx\, dx = \frac{1}{2}\left[\frac{\sin(m+n)x}{m+n} + \frac{\sin(m-n)x}{m-n}\right]_{-\pi}^{\pi} = 0$$

при $m \neq n$.

Интеграл III обращается в нуль на отрезке $[-\pi, \pi]$ в силу нечетности подынтегральной функции. Полагая во втором и третьем интегралах $m = 0$, получаем

$$\int\limits_{-\pi}^{\pi} 1 \cdot \cos nx\, dx = 0, \qquad \int\limits_{-\pi}^{\pi} 1 \cdot \sin nx\, dx = 0.$$

Итак, тригонометрическая система (1) ортогональна на отрезке $[-\pi, \pi]$, а следовательно, и на любом отрезке $[a, 2\pi + a]$.

Вычислим нормы функций, входящих в систему (1). Имеем:

$$\|1\| = \sqrt{\int\limits_{-\pi}^{\pi} 1^2 dx} = \sqrt{2\pi}, \qquad \|\sin nx\| = \sqrt{\int\limits_{-\pi}^{\pi} \sin^2 nx\, dx} = \sqrt{\pi}, \qquad (2)$$

$$\|\cos nx\| = \sqrt{\int\limits_{-\pi}^{\pi} \cos^2 nx\, dx} = \sqrt{\pi} \quad (n = 1,\ 2,\ \ldots).$$

Пусть дана непрерывная периодическая функция с периодом 2π. Составим *тригонометрический полином*

$$Q_m(x) = \frac{a_0}{2} + \sum_{k=1}^{m} (a_k \cos kx + b_k \sin kx); \tag{3}$$

слагаемые $u_0 = a_0/2$, $u_k = a_k \cos kx + b_k \sin kx$ $(k = 1, 2, \ldots)$ обычно называются *гармониками*.

Для того чтобы квадратичное отклонение полинома $Q_m(x)$ от функции $f(x)$ было минимальным, коэффициенты a_0, a_k, b_k должны быть коэффициентами Фурье функции $f(x)$ относительно тригонометрической системы. Отсюда на основании § 8 (формула (7)), учитывая соотношения (2), получаем

$$a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos kx\, dx, \quad b_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin kx\, dx \tag{4}$$

$$(k = 0, 1, 2, \ldots, m).$$

Коэффициенты a_k, b_k называются *тригонометрическими коэффициентами Фурье* функции $f(x)$, а соответствующий тригонометрический полином (3) — *тригонометрическим полиномом Фурье*. Свободный член записывают в виде $a_0/2$ для того, чтобы коэффициент a_0 получался из первой формулы (4) при $k = 0$.

Из формул (4) следует: если функция $f(x)$ четная, то коэффициенты

$$\left.\begin{aligned} &b_k = 0 \quad (k = 1, 2, \ldots, m), \\ &a_k = \frac{2}{\pi} \int_0^{\pi} f(x) \cos kx\, dx \quad (k = 0, 1, 2, \ldots, m), \end{aligned}\right\} \tag{5}$$

причем

$$Q_m(x) = \frac{a_0}{2} + \sum_{k=1}^{m} a_k \cos kx.$$

Если же $f(x)$ — нечетная функция, то

$$\left.\begin{aligned} &a_k = 0 \quad (k = 0, 1, 2, \ldots, m), \\ &b_k = \frac{2}{\pi} \int_0^{\pi} f(x) \sin kx\, dx \quad (k = 1, 2, \ldots, m), \end{aligned}\right\} \tag{6}$$

причем

$$Q_m(x) = \sum_{k=1}^{m} b_k \sin kx.$$

Подставляя в формулу (3) коэффициенты (4), при $m \to \infty$ получаем для функции $f(x)$ ее *тригонометрический ряд Фурье*

$$\frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx).$$

Представление функции тригонометрическим полиномом Фурье или тригонометрическим рядом Фурье называется *гармоническим анализом*. В простейших случаях коэффициенты тригонометрического полинома Фурье данной функции непосредственно вычисляются по формулам (4).

Если же вычисление интегралов (4) громоздко или функция $f(x)$ задана таблично, то для вычисления коэффициентов Фурье имеются различные приближенные способы. Один из этих способов состоит в том, что интегралы (4) вычисляются приближенно по формуле трапеций (см. [2]).

Рассмотрим отдельно три случая.

Случай 1. Функция $f(x)$ — четная. Применяя формулу трапеций

$$\int_a^b \varphi(x)\,dx \approx \frac{b-a}{n}\left[\frac{1}{2}\varphi(x_0) + \varphi(x_1) + \ldots + \varphi(x_{n-1}) + \frac{1}{2}\varphi(x_n)\right]$$

к интегралам (5), получаем

$$a_k = \frac{2}{\pi}\cdot\frac{\pi}{n}\left[\frac{1}{2}f(x_0)\cos kx_0 + f(x_1)\cos kx_1 + \ldots\right.$$
$$\left.\ldots + f(x_{n-1})\cos kx_{n-1} + \frac{1}{2}f(x_n)\cos kx_n\right],$$

где $x_i = \pi i/n$ $(i = 0, 1, 2, \ldots, n)$.

Введя обозначения:

$$y_0 = \frac{1}{2}f(x_0), \quad y_1 = f(x_1), \quad \ldots, \quad y_{n-1} = f(x_{n-1}), \quad y_n = \frac{1}{2}f(x_n),$$

получим

$$a_k = \frac{2}{n}(y_0 \cos kx_0 + y_1 \cos kx_1 + \ldots + y_n \cos kx_n) = \tag{7}$$
$$= \frac{2}{n}\sum_{i=0}^{n} y_i \cos kx_i \qquad (k = 0, 1, 2, \ldots, m).$$

Пример 1. Построить тригонометрический полином Фурье для четной функции, заданной следующими данными:

x_i	0	$\frac{\pi}{12}$	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$	$\frac{5\pi}{12}$	$\frac{\pi}{2}$	$\frac{7\pi}{12}$	$\frac{2\pi}{3}$	$\frac{3\pi}{4}$	$\frac{5\pi}{6}$	$\frac{11\pi}{12}$	π
$f(x_i)$	9,55	9,46	9,25	8,96	8,58	8,10	7,59	7,00	6,34	5,56	4,80	4,21	4,00

Решение. В формуле (7) берем $n=12$, иными словами, отрезок $[-\pi, \pi]$ делим на 24 части. Ограничимся шестью гармониками, не считая нулевой, т. е. примем $m=6$. Все вычисления для отыскания коэффициентов a_k располагаем по схеме А (таблица 8).

В строке $\sum$ записываем сумму $\sum_{i=0}^{n} y_i \cos kx = s_k$ для $k=0, 1, 2, \ldots, 6$. Разделив s_k на $\frac{n}{2}=6$, получаем приближенные значения коэффициентов a_k, приведенные в последней строке схемы А.

Таблица 8

Схема А. Вычисление коэффициентов Фурье для случая четной функции

i	y_i	$\cos x_i$	$\cos 2x_i$	$\cos 3x_i$	$\cos 4x_i$	$\cos 5x_i$	$\cos 6x_i$
0	4,78	1	1	1	1	1	1
1	9,46	0,96593	0,86603	0,70711	0,5	0,25882	0
2	9,25	0,86603	0,5	0	—0,5	—0,86603	—1
3	8,96	0,70711	0	—0,70711	—1	—0,70711	0
4	8,58	0,5	—0,5	—1	—0,5	0,5	1
5	8,10	0,25882	—0,86603	—0,70711	0,5	0,96593	0
6	7,59	0	—1	0	1	0	—1
7	7,00	—0,25882	—0,86603	0,70711	0,5	—0,96593	0
8	6,34	—0,5	—0,5	1	—0,5	—0,5	1
9	5,56	—0,70711	0	0,70711	—1	0,70711	0
10	4,80	—0,86603	0,5	0	—0,5	0,86603	—1
11	4,21	—0,96593	0,86603	—0,70711	0,5	—0,25882	0
12	2,00	—1	1	—1	1	—1	1
$\sum$	86,63	15,52	—2,48	1,06	—0,25	0	0,06
a_k	14,44	2,59	—0,41	0,18	—0,04	0	0,01

Искомый полином имеет вид

$$Q_6(x) = 7{,}22 + 2{,}59 \cos x - 0{,}41 \cos 2x + 0{,}18 \cos 3x - 0{,}04 \cos 4x + 0{,}01 \cos 6x.$$

Для проверки вычислим:

$$Q_6(0) = 7{,}22 + 2{,}59 - 0{,}41 + 0{,}18 - 0{,}04 + 0{,}01 = 9{,}55.$$

Полученный результат хорошо согласуется с заданной таблицей.

Случай 2. Функция $f(x)$ — нечетная. В этом случае задача нахождения тригонометрического полинома сводится к определению коэффициентов b_k (формула (6)):

$$b_k = \frac{2}{\pi} \int_0^{\pi} f(x) \sin kx \, dx.$$

Применяя к этому интегралу формулу трапеций, получим

$$b_k=\frac{2}{\pi}\cdot\frac{\pi}{n}\left[\frac{1}{2}f(x_0)\sin kx_0+\right.$$
$$\left.+f(x_1)\sin kx_1+\ldots+f(x_{n-1})\sin kx_{n-1}+\frac{1}{2}f(x_n)\sin kx_n\right],$$

где $x_i=\pi i/n$ $(i=0,\ 1,\ \ldots,\ n)$.

Введя обозначения $y_i=f(x_i)$ $(i=0,\ 1,\ \ldots,\ n)$ и замечая, что $\sin kx_0=\sin kx_m=0$, получаем

$$b_k=\frac{2}{n}(y_1\sin kx_1+y_2\sin kx_2+\ldots+y_{n-1}\sin kx_{n-1})=$$
$$=\frac{2}{n}\sum_{i=1}^{n-1}y_i\sin kx_i\quad(k=1,\ 2,\ \ldots,\ m).$$

Пример 2. Построить тригонометрический полином Фурье $Q_6(x)$ для нечетной функции $f(x)$, заданной табличными данными:

x_i	0	$\frac{\pi}{12}$	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$	$\frac{5\pi}{12}$	$\frac{\pi}{2}$	$\frac{7\pi}{12}$	$\frac{2\pi}{3}$	$\frac{3\pi}{4}$	$\frac{5\pi}{6}$	$\frac{11\pi}{12}$	π
$f(x_i)$	0	—0,07	—0,10	—0,11	0,16	0,65	1,19	1,55	1,49	1,11	0,58	0,13	0

Решение. Все вычисления располагаем по схеме Б (таблица 8а), аналогичной схеме А для четной функции.

Таблица 8а

Схема Б. Вычисление коэффициентов Фурье для случая нечетной функции

i	y_i	$\sin x_i$	$\sin 2x_i$	$\sin 3x_i$	$\sin 4x_i$	$\sin 5x_i$	$\sin 6x_i$
1	—0,07	0,25882	0,00000	0,70711	0,86603	0,96593	1
2	—0,10	0,50000	0,86603	1	0,86603	0,50000	0
3	—0,11	0,70711	1	0,70711	0	—0,70711	—1
4	0,16	0,86603	0,86603	0	—0,86603	—0,86603	0
5	0,65	0,96593	0,50000	—0,70711	—0,86603	0,25882	1
6	1,19	1	0	—1	0	1	0
7	1,55	0,96593	—0,50000	—0,70711	0,86603	0,25882	—1
8	1,49	0,86603	—0,86603	0	0,86603	—0,86603	0
9	1,11	0,70711	—1	0,70711	0	—0,70711	1
10	0,58	0,50000	—0,86603	1	—0,86603	0,50000	0
11	0,13	0,25882	—0,50000	0,70711	—0,86603	0,96593	—1
$\sum$	—	5,7066	—0,178	—1,5161	1,1691	—0,338	0,120
b_k	—	0,952	—0,030	—0,252	0,194	—0,056	0,020

В последней строке схемы Б записаны приближенные значения коэффициентов b_k.

Искомый полином имеет вид

$$Q_6(x) = 0{,}952 \sin x - 0{,}030 \sin 2x - 0{,}252 \sin 3x + 0{,}194 \sin 4x - \\ - 0{,}056 \sin 5x + 0{,}020 \sin 6x.$$

Для проверки вычислим $Q_6(\pi/2)$. Имеем

$$Q_6(\pi/2) = 0{,}952 + 0{,}252 - 0{,}056 = 1{,}148$$

вместо $f(\pi/2) = 1{,}19$.

Случай 3. Функция $f(x)$ не обладает четностью. Очевидно,

$$f(x) = \frac{1}{2}[f(x) + f(-x)] + \frac{1}{2}[f(x) - f(-x)],$$

причем первый член правой части четен, а второй — нечетен. Обозначая

$$F(x) = \frac{1}{2}[f(x) + f(-x)] \text{ и } G(x) = \frac{1}{2}[f(x) - f(-x)],$$

где $F(x)$ — четная, а $G(x)$ — нечетная функции, представим всякую функцию $f(x)$ в виде суммы четной и нечетной функций:

$$f(x) = F(x) + G(x).$$

Тригонометрический полином Фурье функции $f(x)$ складывается из тригонометрических полиномов Фурье функций $F(x)$ и $G(x)$. Но тригонометрический полином функции $F(x)$, содержащий лишь косинусы кратных дуг, находится по схеме А, а для функции $G(x)$ — по схеме Б. Таким образом, общий случай сводится к разобранным выше.

Пример 3. Построить тригонометрический полином Фурье функции $f(x)$, заданной следующими табличными данными:

x_i	π	$-\frac{11\pi}{12}$	$-\frac{5\pi}{6}$	$-\frac{3\pi}{4}$	$-\frac{2\pi}{3}$	$-\frac{7\pi}{12}$	$-\frac{\pi}{2}$	$-\frac{5\pi}{12}$
$f(x_i)$	0,04328	0,05613	0,07353	0,09536	0,1237	0,1604	0,2080	0,2698

x_i	$-\frac{\pi}{3}$	$-\frac{\pi}{4}$	$-\frac{\pi}{6}$	$-\frac{\pi}{12}$	0	$\frac{\pi}{12}$	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$
$f(x_i)$	0,3535	0,4584	0,5945	0,7711	1,0000	1,2969	1,6820	2,1815	2,8292

x_i	$\frac{5\pi}{12}$	$\frac{\pi}{2}$	$\frac{7\pi}{12}$	$\frac{2\pi}{3}$	$\frac{3\pi}{4}$	$\frac{5\pi}{6}$	$\frac{11\pi}{12}$	π
$f(x_i)$	3,7062	4,8066	6,2399	8,0849	10,4860	13,5993	17,8142	23,1040

Решение. Определяем значения функций

$$F(x)=\frac{1}{2}[f(x)+f(-x)] \text{ и } G(x)=\frac{1}{2}[f(x)-f(-x)]$$

и результаты записываем в таблицах 9 и 9а соответственно.

Таблица 9

Значения функции $F(x)$

x_i	0	$\frac{\pi}{12}$	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$	$\frac{5\pi}{12}$	$\frac{\pi}{2}$
y_i	0,6000	1,0340	1,13825	1,31095	1,59135	1,9880	2,5073

x_i	$\frac{7\pi}{12}$	$\frac{2\pi}{3}$	$\frac{3\pi}{4}$	$\frac{5\pi}{6}$	$\frac{11\pi}{12}$	π
y_i	3,20015	4,1043	5,29068	6,83642	8,93516	5,78682

$$\left(y_0=\frac{1}{2}f(x_0);\quad y_n=\frac{1}{2}f(x_n)\right).$$

Таблица 9а

Значения функции $G(x)$

x_i	0	$\frac{\pi}{12}$	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$	$\frac{5\pi}{12}$	$\frac{\pi}{2}$
y_i	0,0000	0,2629	0,54735	0,86155	1,23785	1,7182	2,2993

x_i	$\frac{7\pi}{12}$	$\frac{2\pi}{3}$	$\frac{3\pi}{4}$	$\frac{5\pi}{6}$	$\frac{11\pi}{12}$	π
y_i	3,03975	3,9806	5,19532	6,76288	8,87904	11,53036

Вычисление коэффициентов функции $F(x)$ по схеме А дает

$$\frac{1}{2}a_0=3{,}686;\quad a_1=-3{,}705;\quad a_2=1{,}510;\quad a_3=-0{,}776;$$
$$a_4=0{,}488;\quad a_5=-0{,}366;\quad a_6=0{,}250.$$

Таким образом,

$$F_6(x)=3{,}686-3{,}705\cos x+1{,}510\cos 2x-0{,}776\cos 3x+$$
$$+0{,}488\cos 4x-0{,}366\cos 5x+0{,}250\cos 6x. \qquad (8)$$

Для функции $G(x)$ используем схему Б и получаем:

$$b_1=3{,}475;\qquad b_2=-2{,}699;\quad b_3=2{,}066;$$
$$b_4=-1{,}698;\quad b_5=1{,}347;\quad b_6=-0{,}934.$$

Следовательно,

$$G_6(x)=3{,}475\sin x-2{,}699\sin 2x+2{,}066\sin 3x-1{,}698\sin 4x+$$
$$+1{,}347\sin 5x-0{,}934\sin 6x. \qquad (9)$$

Складывая полиномы (8) и (9), получаем искомый тригонометрический полином

$$Q_6(x)=3{,}686-3{,}705\cos x+3{,}475\sin x+1{,}510\cos 2x-$$
$$-2{,}699\sin 2x-0{,}776\cos 3x+2{,}066\sin 3x+0{,}488\cos 4x-$$
$$-1{,}698\sin 4x-0{,}366\cos 5x+1{,}347\sin 5x+0{,}250\cos 6x-0{,}934\sin 6x.$$

Для проверки вычислим $Q_6(\pi/2)$. Имеем

$$Q_6(\pi/2)=3{,}686+3{,}475-1{,}510-2{,}066+0{,}488+1{,}347-0{,}250=5{,}170$$

вместо $f(\pi/2)=4{,}8066$.

Относительная погрешность приближения $\approx 5\%$.

§ 10. Полиномы Лежандра

Полиномы Лежандра определяются следующей формулой Родрига:

$$P_n(x)=\frac{1}{2^n n!}\frac{d^n}{dx^n}(x^2-1)^n \quad (n=0,\ 1,\ 2,\ \ldots). \qquad (1)$$

В частности, имеем:

$$\left.\begin{aligned} &P_0(x)=1, && P_3(x)=\frac{1}{2}(5x^3-3x),\\ &P_1(x)=x, && \\ &P_2(x)=\frac{1}{2}(3x^2-1), && P_4(x)=\frac{1}{8}(35x^4-30x^2+3)\end{aligned}\right\} \qquad (2)$$

и т. д. (см. приложение II). Графики этих полиномов для $n=0$, 1, 2, 3 и 4. приведены на рис. 4. Из формулы (1) видно, что $P_n(x)$ являются четными функциями при $n=2m$ и нечетными — при $n=2m+1$; причем $P_n(1)=1$ и $P_n(-1)=(-1)^n$.

Теорема 1. *Полиномы Лежандра образуют ортогональную систему на отрезке* $[-1, 1]$, *т. е.*

$$\int_{-1}^{1} P_n(x) P_m(x)\, dx = 0 \text{ при } m \neq n \quad (m,\ n = 0,\ 1,\ 2,\ \ldots).$$

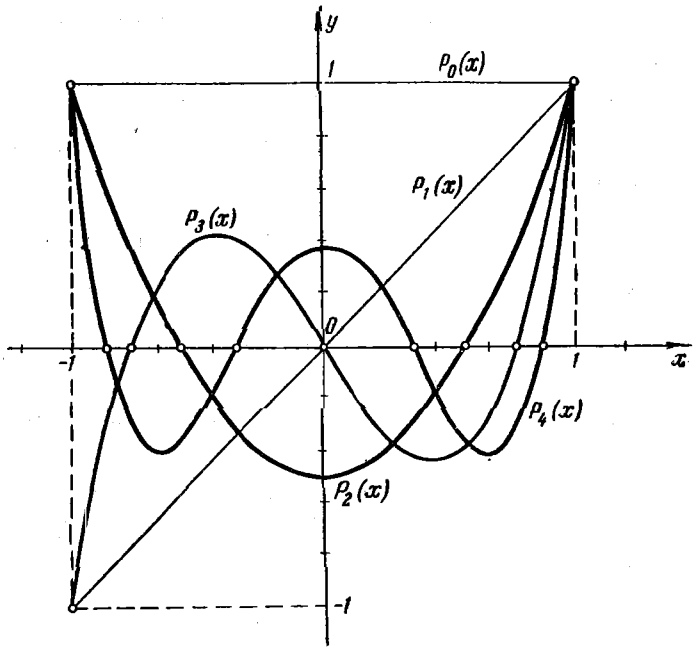

Рис. 4.

Доказательство. Имеем

$$I_{mn} = \int_{-1}^{1} P_n(x) P_m(x)\, dx =$$

$$= \frac{1}{2^{m+n} m!\, n!} \int_{-1}^{1} \frac{d^m}{dx^m} (x^2-1)^m \frac{d^n}{dx^n} (x^2-1)^n\, dx. \tag{3}$$

Пусть $m < n$. Применяя к интегралу, стоящему в правой части равенства (3), формулу интегрирования по частям и учитывая, что

$$\left.\frac{d^k (x^2-1)^n}{dx^k}\right|_{x=-1}^{x=1} = 0 \quad \text{при } k < n \tag{4}$$

(так как значения $x=\pm 1$ являются нулями кратности n для функции $(x^2-1)^n$), получим

$$I_{mn}=\frac{1}{2^{m+n}n!m!}\left\{\frac{d^m}{dx^m}(x^2-1)^m\frac{d^{n-1}}{dx^{n-1}}(x^2-1)^n\Big|_{-1}^{1}-\right.$$

$$\left.-\int_{-1}^{1}\frac{d^{m+1}}{dx^{m+1}}(x^2-1)^m\frac{d^{n-1}}{dx^{n-1}}(x^2-1)^n\,dx\right\}=$$

$$=\frac{-1}{2^{m+n}n!m!}\int_{-1}^{1}\frac{d^{m+1}}{dx^{m+1}}(x^2-1)^m\frac{d^{n-1}}{dx^{n-1}}(x^2-1)^n\,dx.$$

После n-кратного интегрирования по частям формулы (3) в силу соотношения (4) будем иметь

$$I_{mn}=\frac{(-1)^n}{2^{m+n}n!m!}\int_{-1}^{1}\frac{d^{m+n}}{dx^{m+n}}(x^2-1)^m(x^2-1)^n\,dx. \qquad (5)$$

Но так как $m<n$, то, очевидно, $\frac{d^{m+n}}{dx^{m+n}}(x^2-1)^m\equiv 0$ и, следовательно,

$$I_{mn}=\int_{-1}^{1}P_m(x)P_n(x)\,dx=0 \quad \text{при } m\neq n.$$

Замечание 1. Подсчитаем норму полинома $P_n(x)$. Полагая $m=n$ в формуле (5) и принимая во внимание, что

$$\frac{d^{2n}}{dx^{2n}}(x^2-1)^n=\frac{d^{2n}}{dx^{2n}}(x^{2n}-C_n^1x^{2n-2}+\ldots)=(2n)!,$$

получим

$$I_{nn}=\int_{-1}^{1}P_n^2(x)\,dx=\frac{(-1)^n}{2^{2n}(n!)^2}\int_{-1}^{1}\frac{d^{2n}}{dx_{2n}}(x^2-1)^n(x^2-1)^n\,dx=$$

$$=\frac{(-1)^n(2n)!}{2^{2n}(n!)^2}\int_{-1}^{1}(-1)^n(1-x^2)^n\,dx=\frac{(2n)!}{2^{2n}(n!)^2}\int_{-1}^{1}(1-x)^n(1+x)^n dx. \qquad (6)$$

Последний интеграл вычисляется при помощи n-кратного интегрирования по частям:

$$\int_{-1}^{1}(1-x)^n(1+x)^n\,dx=$$

$$=(1-x)^n\frac{(1+x)^{n+1}}{n+1}\Big|_{-1}^{1}+\frac{n}{n+1}\int_{-1}^{1}(1-x)^{n-1}(1+x)^{n+1}\,dx=$$

$$=\frac{n(n-1)\ldots 1}{(n+1)(n+2)\ldots 2n}\int_{-1}^{1}(1+x)^{2n}\,dx=\frac{(n!)^2}{(2n)!}\cdot\frac{2^{2n+1}}{2n+1}. \qquad (7)$$

Подставляя полученный результат (7) в формулу (6), будем иметь

$$\int_{-1}^{1} P_n^2(x)\,dx = \frac{(2n)!}{2^{2n}(n!)^2}\cdot\frac{(n!)^2}{(2n)!}\cdot\frac{2^{2n+1}}{2n+1} = \frac{2}{2n+1} \quad (n=0,\ 1,\ 2,\ \ldots). \tag{8}$$

Таким образом,

$$\|P_n(x)\|^2 = \frac{2}{2n+1} \quad \text{или} \quad \|P_n(x)\| = \sqrt{\frac{2}{2n+1}}.$$

Нормированные полиномы Лежандра

$$\hat{P}_n(x) = \sqrt{\frac{2n+1}{2}}\,P_n(x) \quad (n=0,\ 1,\ 2,\ \ldots)$$

образуют на отрезке $[-1,\ 1]$ ортонормированную систему полиномов, для которых

$$\int_{-1}^{1} \hat{P}_n(x)\hat{P}_m(x)\,dx = \begin{cases} 0 & \text{при } m \neq n, \\ 1 & \text{при } m = n. \end{cases}$$

Замечание 2. Полиномы Лежандра обладают усиленным свойством ортогональности

$$\int_{-1}^{1} P_m(x)\,R_k(x)\,dx = 0 \quad (k < m), \tag{9}$$

где $R_k(x)$ — любой полином степени k, меньшей m.

В самом деле, очевидно, справедливо соотношение

$$R_k(x) = c_0P_k(x) + c_1P_{k-1}(x) + \ldots + c_kP_0(x),$$

коэффициенты которого $c_0,\ c_1,\ \ldots,\ c_k$ могут быть определены с помощью последовательных делений. Отсюда

$$\int_{-1}^{1} P_m(x)\,R_k(x)\,dx = \int_{-1}^{1} P_m(x)\sum_{i=0}^{k} c_iP_{k-i}(x)\,dx =$$

$$= \sum_{i=0}^{k} c_i \int_{-1}^{1} P_m(x)\,P_{k-i}(x)\,dx = 0.$$

Замечание 3. С помощью линейного преобразования

$$z = \frac{b-a}{2}x + \frac{b+a}{2},$$

где $-1 \leqslant x \leqslant 1$, можно получить полиномы

$$\tilde{P}_m(z) = P_m\left(\frac{z - \frac{b+a}{2}}{\frac{b-a}{2}}\right),$$

ортогональные на отрезке $[a, b]$, т. е. такие, что

$$\int_a^b \tilde{P}_m(z)\, \tilde{P}_n(z)\, dz = 0 \quad \text{при } m \neq n.$$

Используя формулу (8), легко показать, что

$$\int_a^b \tilde{P}_n^2(z)\, dz = \frac{b-a}{2n+1} \qquad (n = 0, 1, 2, \ldots).$$

Теорема 2. *Полином Лежандра $P_n(x)$ при $n \geqslant 1$ имеет n различных действительных корней, которые расположены на интервале* $(-1, 1)$.

Доказательство. Пусть $r_1, r_2, \ldots, r_k$ $(r_1 < r_2 < \ldots < r_k)$ — действительные корни полинома $P_n(x)$, имеющие нечетную кратность и лежащие на интервале $(-1, 1)$. Очевидно, что $k \leqslant n$. Произведение $(x - r_1)(x - r_2) \ldots (x - r_k) P_n(x)$ не меняет знака на интервале $(-1, 1)$, так как все его действительные корни на этом интервале имеют четную кратность. Поэтому

$$\int_{-1}^{1} (x - r_1)(x - r_2) \ldots (x - r_k) P_n(x)\, dx \neq 0.$$

Отсюда, согласно замечанию 2 (см. формулу (9)), следует, что степень полинома $R_k(x) = (x - r_1)(x - r_2) \ldots (x - r_k)$ не может быть ниже n, т. е. $k \geqslant n$. Но так как $k \leqslant n$, то получаем, что $k = n$, т. е. все n корней полинома $P_n(x)$ действительные, простые и лежат на интервале $(-1, 1)$.

Замечание. Полиномы Лежандра можно определить также при помощи так называемой *производящей функции* [1]

$$H(x, r) \equiv \frac{1}{(1 - 2xr + r^2)^{1/2}} = P_0(x) + P_1(x) r + \\ + P_2(x) r^2 + \ldots + P_n(x) r^n + \ldots \quad (|r| \leqslant 1). \quad (10)$$

Формула (10) играет важную роль в теории потенциала.

В частности, полагая $x = 1$ в формуле (10), получим

$$\frac{1}{1-r} = P_0(1) + P_1(1) r + \ldots + P_n(1) r^n + \ldots$$

Но так как

$$\frac{1}{1-r}=1+r+r^2+\ldots,$$

то $P_n(1)=1$ $(n=0, 1, 2, \ldots)$. Аналогично, положив $x=-1$, получаем $P_n(-1)=(-1)^n$ $(n=0, 1, 2, \ldots)$.

Из разложения (10) легко получить рекуррентную формулу, связывающую три последовательных полинома Лежандра. Действительно, дифференцируя $H(x, r)$ по r, имеем

$$\frac{\partial H}{\partial r}=\frac{x-r}{(1-2xr+r^2)^{3/2}}=\frac{x-r}{1-2xr+r^2}H.$$

Отсюда, учитывая, что

$$\frac{\partial H}{\partial r}=P_1(x)+2P_2(x)r+\ldots+nP_n(x)r^{n-1}+\ldots,$$

имеем

$$(1-2xr+r^2)[P_1(x)+2P_2(x)r+\ldots+nP_n(x)r^{n-1}+\ldots]-$$
$$-(x-r)[P_0(x)+P_1(x)r+\ldots+P_n(x)r^n+\ldots]\equiv 0.$$

Собирая все члены, содержащие r^n, и приравнивая нулю полученный коэффициент при r^n, находим нужный результат

$$(n+1)P_{n+1}(x)-x(2n+1)P_n(x)+nP_{n-1}(x)=0 \qquad (11)$$
$$(n=1, 2, \ldots).$$

С помощью формулы (11) удобно находить последовательные полиномы Лежандра. Например, полагая $n=4$, будем иметь

$$P_5(x)=\frac{1}{5}[9x\,P_4(x)-4P_3(x)]=$$
$$=\frac{1}{5}\left[\frac{9}{8}(35x^5-30x^3+3x)-(10x^3-6x)\right]=$$
$$=\frac{1}{40}(315x^5-350x^3+75x)=\frac{1}{8}(63x^5-70x^3+15x).$$

Сводка формул для полиномов Лежандра и их обобщений, а также таблицы и графики имеются в книге Янке и Эмде [4].

Пусть на отрезке $[-1, 1]$ задана функция $f(x)$. Приблизим эту функцию линейной комбинацией из полиномов Лежандра

$$f(x)\approx c_0P_0(x)+c_1P_1(x)+\ldots+c_mP_m(x),$$

причем коэффициенты $c_0, c_1, \ldots, c_m$ подберем так, чтобы величина отклонения

$$S_m=\int_{-1}^{1}\left[f(x)-\sum_{k=0}^{m}c_kP_k(x)\right]^2dx$$

была минимальная.

В § 8 доказано, что в этом случае c_k являются коэффициентами Фурье относительно полиномов $P_k(x)$, т. е.

$$c_k = \frac{1}{\|P_k(x)\|^2} \int_{-1}^{1} f(x) P_k(x)\, dx = \frac{2k+1}{2} \int_{-1}^{1} f(x) P_k(x)\, dx. \qquad (12)$$

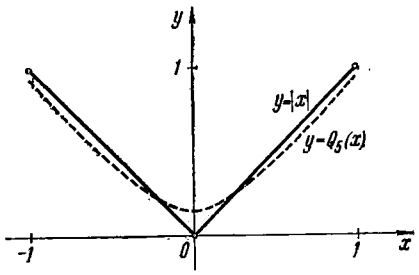

Рис. 5.

Из § 8 (формула (10)) следует также

$$S_m = \int_{-1}^{1} f^2(x)\, dx - \sum_{k=0}^{m} \frac{2}{2k+1} c_k^2. \qquad (13)$$

Пример 1. Функцию $f(x) = |x|$ на отрезке $[-1, 1]$ квадратично аппроксимировать полиномом пятой степени.

Решение. Полином $Q_5(x)$ ищем в следующем виде:

$$Q_5(x) = c_0P_0(x) + c_1P_1(x) + c_2P_2(x) + c_3P_3(x) + c_4P_4(x) + c_5P_5(x), \qquad (14)$$

где $P_k(x)$ $(k=0, 1, \ldots)$ — полиномы Лежандра. Так как функция $|x|$ четная и $P_k(x)$ четны при k четном и нечетны при k нечетном, то из формулы (12) получаем

$$c_{2k} = (4k+1) \int_0^1 xP_{2k}(x)\, dx \text{ и } c_{2k+1} = 0 \quad (k=0, 1, 2).$$

Отсюда, используя формулы (2), находим:

$$c_0 = \int_0^1 x\, dx = \frac{1}{2}, \qquad c_2 = \frac{5}{2} \int_0^1 x(3x^2-1)\, dx = \frac{5}{8},$$

$$c_4 = \frac{9}{8} \int_0^1 x(35x^4 - 30x^2 + 3)\, dx = -\frac{3}{16}$$

и

$$c_1 = c_3 = c_5 = 0.$$

Подставляя эти значения коэффициентов в формулу (14), получим

$$Q_5(x) = \frac{1}{2} + \frac{5}{16}(3x^2-1) - \frac{3}{128}(35x^4 - 30x^2 + 3) = \frac{15}{128}(-7x^4 + 14x^2 + 1).$$

Следовательно (рис. 5),

$$|x| \approx \frac{15}{128}(1 + 14x^2 - 7x^4) \quad \text{при } |x| \leqslant 1.$$

Пример 2. Функцию $f(x) = 3^x$ на отрезке $[-1, 1]$ квадратично аппроксимировать полиномом третьей степени (см. пример 2 из § 7).

Решение. Полином ищем в виде

$$Q_3(x)=c_0P_0(x)+c_1P_1(x)+c_2P_2(x)+c_3P_3(x),$$

где $P_k(x)$ $(k=0,\ 1,\ 2,\ 3)$ — полиномы Лежандра. Согласно формулам (12) получаем:

$$c_0=\frac{1}{2}\int_{-1}^{1}3^x dx=1{,}2137,$$

$$c_1=\frac{3}{2}\int_{-1}^{1}x\cdot 3^x\,dx=1{,}2371,$$

$$c_2=\frac{5}{2}\int_{-1}^{1}(3x^2-1)\cdot\frac{1}{2}\cdot 3^x\,dx=0{,}4384,$$

$$c_3=\frac{7}{2}\int_{-1}^{1}\frac{1}{2}(5x^3-3x)\cdot 3^x dx=0{,}09345.$$

Следовательно,

$$Q_3(x)=1{,}2137+1{,}2371x+\frac{0{,}4384}{2}(3x^2-1)+$$
$$+\frac{0{,}09345}{2}(5x^3-3x)=0{,}2336x^3+0{,}6576x^2+1{,}0968x+0{,}9945.$$

Как и следовало ожидать, мы получили подтверждение результата, найденного обычным способом наименьших квадратов, без применения ортогональных полиномов (см. § 7, пример 2).

§ 11. Ортогональность с весом

Пусть на отрезке $[a,\ b]$ задана положительная непрерывная функция $\rho(x)$.

Определение. Система функций $\{\varphi_n(x)\}$, заданная на отрезке $[a,\ b]$, называется *ортогональной* на этом отрезке с *весом* $\rho(x)$, если

$$\int_a^b \rho(x)\,\varphi_k(x)\,\varphi_m(x)\,dx=0 \quad \text{при} \quad k\neq m. \tag{1}$$

Из ортогональности системы $\{\varphi_n(x)\}$ с весом $\rho(x)$ следует обычная ортогональность системы $\psi_n(x)=\sqrt{\rho(x)}\,\varphi_n(x)$ $(n=0,1,2,\ldots)$. Действительно,

$$\int_a^b \psi_k(x)\,\psi_m(x)\,dx=\int_a^b \rho(x)\,\varphi_k(x)\,\varphi_m(x)\,dx=0 \quad \text{при} \quad k\neq m,$$

а это и есть условие ортогональности.

Пусть на отрезке $[a, b]$ задана непрерывная функция $f(x)$ и система функций $\varphi_0(x), \varphi_1(x), \ldots, \varphi_m(x), \ldots$ ортогональна на этом отрезке с весом $\rho(x)$. Составим обобщенный полином

$$Q_m(x) = c_0\varphi_0(x) + c_1\varphi_1(x) + \ldots + c_m\varphi_m(x)$$

и подберем коэффициенты $c_0, c_1, \ldots, c_n$ так, чтобы отклонение, определяемое формулой

$$I_m = \int_a^b \rho(x)\,[f(x) - Q_m(x)]^2\,dx, \tag{2}$$

было минимально при фиксированном m.

Формулу (2) можно записать в следующем виде:

$$I_m = \int_a^b \left[f(x)\sqrt{\rho(x)} - \sum_{k=0}^{m} c_k\sqrt{\rho(x)}\,\varphi_k(x)\right]^2 dx.$$

Отсюда вытекает, что наша задача сводится к разобранной выше задаче (§ 8) о наилучшей непрерывной квадратичной аппроксимации функции $f(x)\sqrt{\rho(x)}$ с помощью линейной комбинации ортогональных на отрезке $[a, b]$ функций $\sqrt{\rho(x)}\,\varphi_k(x)$ $(k = 0, 1, 2, \ldots, m)$.

На основании результатов § 8 получаем, что коэффициенты c_k, дающие минимум интегралу I_m, представляют собой коэффициенты Фурье функции $f(x)\sqrt{\rho(x)}$ относительно системы функций $\{\sqrt{\rho(x)}\,\varphi_k(x)\}$ [см. § 8, (6)]:

$$c_k = \frac{\int_a^b \sqrt{\rho(x)}\,f(x)\sqrt{\rho(x)}\,\varphi_k(x)\,dx}{\int_a^b [\sqrt{\rho(x)}\,\varphi_k(x)]^2\,dx} = \frac{\int_a^b \rho(x)\,f(x)\,\varphi_k(x)\,dx}{\int_a^b \rho(x)\,\varphi_k^2(x)\,dx} \tag{3}$$

$$(k = 0, 1, 2, \ldots, m).$$

Величина минимального отклонения I_m определяется по формуле (10) из § 8 и будет равна в нашем случае

$$I_m = \int_a^b \rho(x) f^2(x)\,dx - \sum_{k=0}^{m} c_k^2 \int_a^b \rho(x)\,\varphi_k^2(x)\,dx.$$

Эти формулы можно получить также непосредственно, решая соответствующую задачу на нахождение минимума функции

$$I_m = I_m(c_0, c_1, c_2, \ldots, c_m).$$

В частности, если вес $\rho(x) = 1$, то приходим к прежним формулам.

§ 12. Полиномы Чебышева

В качестве примера системы функций, ортогональной с весом, приведем полиномы Чебышева, которые известны еще и тем, что являются полиномами, наименее уклоняющимися от нуля.

Полиномы Чебышева $T_n(x)$ определяются формулами

$$T_n(x)=\frac{(x+\sqrt{x^2-1})^n+(x-\sqrt{x^2-1})^n}{2^n} \quad (n=1,\ 2,\ \ldots). \tag{1}$$

Функция (1) действительно является полиномом. В самом деле, при возведении в степень первой и второй скобок радикалы сохраняются во всех четных членах, причем в первой скобке с плюсом, а во второй с минусом. Поэтому при сложении они уничтожаются и останутся лишь члены, содержащие x в целых степенях, причем старшая степень равна n.

Запишем три первых полинома Чебышева в развернутом виде:

$$T_0(x)=1,$$
$$T_1(x)=\frac{1}{2}\left[(x+\sqrt{x^2-1})+(x-\sqrt{x^2-1})\right]=x,$$
$$T_2(x)=\frac{1}{2^2}\left[(x+\sqrt{x^2-1})^2+(x-\sqrt{x^2-1})^2\right]=x^2-\frac{1}{2}.$$

Покажем, что коэффициент при старшем члене x^n в полиноме $T_n(x)$ равен единице.

Действительно, старший коэффициент a_n всякого полинома $R_n(x)=a_0+a_1x+\ldots+a_nx^n$ степени n можно определить с помощью предельного перехода

$$a_n=\lim_{x\to\infty}\frac{R_n(x)}{x^n}.$$

В нашем случае

$$\frac{T_n(x)}{x^n}=\frac{\left(1+\sqrt{1-\frac{1}{x^2}}\right)^n+\left(1-\sqrt{1-\frac{1}{x^2}}\right)^n}{2^n},$$

откуда

$$\lim_{x\to\infty}\frac{T_n(x)}{x^n}=1.$$

Приведем другую формулу для полиномов Чебышева. Обычно полиномы Чебышева рассматриваются на отрезке $[-1,\ 1]$. Поэтому можно положить $x=\cos t$, т. е. $t=\arccos x$, где t — новая переменная $(0\leqslant t\leqslant\pi)$. Тогда $\sin t=\sqrt{1-x^2}$ и формула (1) преобразуется к виду

$$T_n(x)=\frac{(\cos t+i\sin t)^n+(\cos t-i\sin t)^n}{2^n} \quad \text{при } n>0.$$

Так как $(\cos t \pm i \sin t)^n = \cos nt \pm i \sin nt$, то окончательно имеем

$$T_n(x) = \frac{1}{2^{n-1}} \cos nt \qquad (n = 1, 2, \ldots) \tag{2}$$

или

$$T_n(x) = \frac{1}{2^{n-1}} \cos(n \arccos x). \tag{3}$$

Заметим, что формулы (2) и (3) неверны при $n = 0$.

Из формулы (2) легко получаются рекуррентные формулы для вычисления полиномов Чебышева при больших n. Так как

$$\cos(n+1)t + \cos(n-1)t = 2 \cos nt \cos t,$$

а согласно (2) $\cos nt = 2^{n-1} T_n(x)$, то

$$2^n T_{n+1}(x) + 2^{n-2} T_{n-1}(x) = 2^n T_n(x)\, x, \tag{4}$$

или

$$T_{n+1}(x) = x T_n(x) - \frac{1}{4} T_{n-1}(x). \tag{5}$$

Полагая, например, $n = 2$ и $n = 3$ в формуле (5), последовательно получаем

$$T_3(x) = x T_2(x) - \frac{1}{4} T_1(x) = x\left(x^2 - \frac{1}{2}\right) - \frac{1}{4} x = x^3 - \frac{3}{4} x,$$

$$T_4(x) = x T_3(x) - \frac{1}{4} T_2(x) = x\left(x^3 - \frac{3}{4} x\right) - \frac{1}{4}\left(x^2 - \frac{1}{2}\right) =$$

$$= x^4 - x^2 + \frac{1}{8}.$$

Тем же способом можно вычислить $T_5(x)$, $T_6(x)$ и т. д.

Первые 12 полиномов Чебышева $T_n(x)$ даны в приложении III.

На рис. 6 приведены графики полиномов Чебышева для $n = 0$, 1, 2, 3.

Докажем некоторые свойства полиномов Чебышева.

Теорема 1. (Свойство ортогональности полиномов Чебышева.) *Полиномы Чебышева $T_n(x)$ образуют на отрезке $[-1, 1]$ ортогональную систему с весом*

$$p(x) = \frac{1}{\sqrt{1-x^2}},$$

т. е.

$$\int_{-1}^{1} T_k(x)\, T_m(x) \frac{dx}{\sqrt{1-x^2}} = 0 \text{ при } m \neq k.$$

Доказательство. При $k > 0$, $m > 0$ и $k \neq m$, полагая $x = \cos t$ и используя формулу (2), имеем

$$\int_{-1}^{1} T_k(x)\, T_m(x) \frac{dx}{\sqrt{1-x^2}} = -\int_{\pi}^{0} \frac{1}{2^{k-1}} \cos kt \frac{1}{2^{m-1}} \cos mt\, dt =$$

$$= \frac{1}{2^{k+m-2}} \int_{-\pi}^{\pi} \cos kt \cos mt\, dt = 0, \qquad (6)$$

так как тригонометрические функции $\{\cos kt\}$ ортогональны на

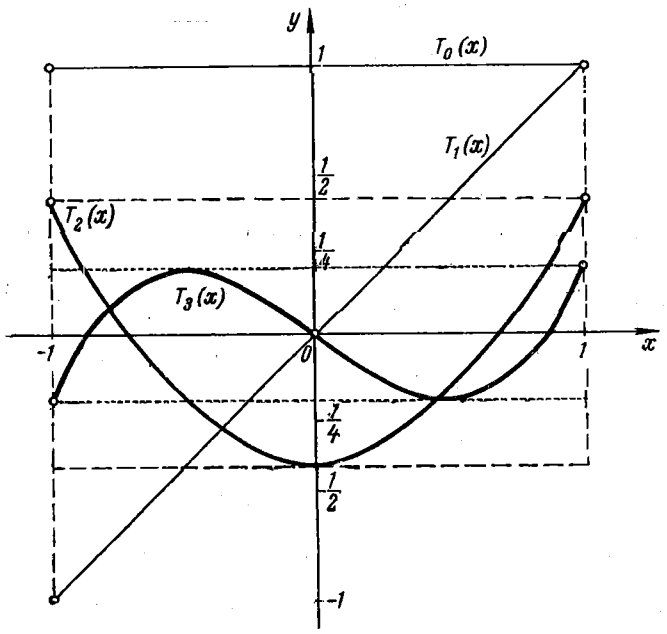

Рис. 6.

отрезке $[-\pi, \pi]$. Легко проверить, что равенство (6) остается справедливым и при $k = 0$, $m \neq 0$.

Замечание. При $m = k \neq 0$ имеем

$$\int_{-1}^{1} T_m^2(x) \frac{dx}{\sqrt{1-x^2}} = \frac{1}{2^{2m-2}} \int_{0}^{\pi} \cos^2 mt\, dt = \frac{1}{2^{2m-2}} \cdot \frac{\pi}{2} = \frac{\pi}{2^{2m-1}}.$$

При $m=k=0$ получаем

$$\int_{-1}^{1} T_0^2(x)\frac{dx}{\sqrt{1-x^2}}=\int_{-1}^{1}\frac{dx}{\sqrt{1-x^2}}=\pi.$$

Теорема 2. (Корни полинома Чебышева.) *Все корни полинома Чебышева ненулевой степени действительны, различны и лежат на интервале* $(-1, 1)$.

Доказательство. Пусть x_k — действительные корни полинома $T_n(x)$, расположенные на отрезке $[-1, 1]$. Полагая $x_k=\cos t_k$ $(0\leqslant t_k\leqslant\pi)$, в силу формулы (2) получим

$$T_n(x_k)=\frac{1}{2^{n-1}}\cos nt_k=0.$$

Отсюда $\cos nt_k=0$, и, следовательно,

$$t_k=\frac{2k-1}{2n}\pi.$$

Таким образом, полином $T_n(x)$ имеет корни

$$x_k=\cos\frac{2k-1}{2n}\pi \qquad (k=1, 2, \ldots, n), \tag{7}$$

которые действительны, различны и лежат на интервале $(-1, 1)$. Так как число этих корней равно степени полинома, то других корней, как действительных, так и комплексных, полином $T_n(x)$ не имеет, т. е. формула (7) дает совокупность всех корней полинома $T_n(x)$.

Теорема 3. (Экстремальное свойство полиномов Чебышева.) *Полином Чебышева* $T_n(x)$ $(n>0)$ *на отрезке* $[-1, 1]$ *имеет* $n+1$ *экстремальных значений, равных между собой по абсолютной величине.*

Доказательство. Имеем

$$T_n(x)=\frac{1}{2^{n-1}}\cos nt, \tag{8}$$

где $x=\cos t$ $(0\leqslant t\leqslant\pi)$. Отсюда заключаем, что $T_n(x)$ достигает экстремума в тех точках, где

$$|\cos nt|=1. \tag{9}$$

Корни уравнения (9), лежащие на отрезке $[0, \pi]$, суть $\tau_k=k\pi/n$ $(k=0, 1, \ldots, n)$. Им соответствуют точки экстремума

$$\xi_k=\cos\frac{k\pi}{n} \qquad (k=0, 1, \ldots, n) \tag{10}$$

полинома Чебышева $T_n(x)$, причем на основании формулы (8) имеем

$$T_n(\xi_k)=\frac{1}{2^{n-1}}\cos k\pi=\frac{(-1)^k}{2^{n-1}}.$$

Следовательно,

$$|T_n(\xi_k)|=\frac{1}{2^{n-1}} \qquad (k=0,\ 1,\ \ldots,n), \tag{11}$$

где ξ_k при k четном суть точки максимума, а при k нечетном — точки минимума полинома $T_n(x)$.

Следствие. Максимальное значение модуля полинома Чебышева $T_n(x)$ $(n>0)$ на отрезке $[-1,\ 1]$ равно $\frac{1}{2^{n-1}}$, т. е.

$$|T_n(x)|\leqslant\frac{1}{2^{n-1}} \quad \text{при} \quad -1\leqslant x\leqslant 1. \tag{12}$$

Замечание. В силу формул (7) и (10) корни и точки экстремума полинома $T_n(x)$ можно построить следующим образом: разделив полуокружность, опирающуюся на отрезок $[-1,\ 1]$ как на диаметр, на $2n$ частей, спроектируем все полученные точки на диаметр (рис. 7, где $n=4$).

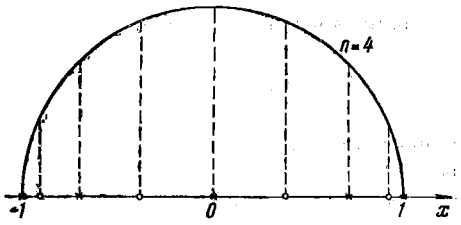

Рис. 7.

Нумеруя проекции справа налево, получим, что все проекции с нечетными индексами являются корнями полинома $T_n(x)$ (на рис. 7 отмечены кружочками), а с четными индексами — его точками экстремума (на рис. 7 отмечены крестиками).

Из геометрических соображений вытекает, что как корни, так и точки экстремума полинома $T_n(x)$ сгущаются к концам отрезка $[-1,\ 1]$.

Пример 1. Найти квадратичное приближение функции $f(x)=|x|$ в промежутке $[-1,\ 1]$ полиномом четвертой степени при весе

$$p(x)=\frac{1}{\sqrt{1-x^2}}.$$

Решение. Полином $Q_4(x)$ будем искать в виде

$$Q_4(x)=c_0T_0(x)+c_1T_1(x)+c_2T_2(x)+c_3T_3(x)+c_4T_4(x), \tag{13}$$

где $T_k(x)$ $(k=0, 1, 2, 3, 4)$ — полиномы Чебышева. Коэффициенты c_k вычисляем по формулам (3) из § 11:

$$c_k=\frac{\int\limits_{-1}^{1}\frac{1}{\sqrt{1-x^2}}\,|x|\,T_k(x)\,dx}{\int\limits_{-1}^{1}\frac{1}{\sqrt{1-x^2}}\,T_k^2(x)\,dx}.$$

Так как $T_{2k+1}(x)$ — функции нечетные, а $T_{2k}(x)$ — четные, то $c_{2k+1}=0$ и

$$c_{2k}=\frac{\int\limits_{0}^{1}\frac{x}{\sqrt{1-x^2}}\,T_{2k}(x)\,dx}{\int\limits_{0}^{1}\frac{1}{\sqrt{1-x^2}}\,T_{2k}(x)\,dx}.$$

Используя для $T_{2k}(x)$ формулу (3), имеем:

$$c_0=\frac{\int\limits_{0}^{1}\frac{x}{\sqrt{1-x^2}}\,dx}{\int\limits_{0}^{1}\frac{1}{\sqrt{1-x^2}}\,dx}=\frac{2}{\pi},$$

$$c_2=\frac{\int\limits_{0}^{1}\frac{x}{\sqrt{1-x^2}}\cdot\frac{1}{2}\cos(2\arccos x)\,dx}{\int\limits_{0}^{1}\frac{1}{\sqrt{1-x^2}}\cdot\frac{1}{4}\cos^2(2\arccos x)\,dx}=\frac{2\int\limits_{0}^{\pi/2}\cos\varphi\cos 2\varphi\,d\varphi}{\int\limits_{0}^{\pi/2}\cos^2 2\varphi\,d\varphi}=\frac{4}{3\pi},$$

$$c_4=\frac{\int\limits_{0}^{1}\frac{x}{\sqrt{1-x^2}}\cdot\frac{1}{2^3}\cos(4\arccos x)\,dx}{\int\limits_{0}^{1}\frac{1}{\sqrt{1-x^2}}\cdot\frac{1}{2^6}\cos^2(4\arccos x)\,dx}=\frac{2^3\int\limits_{0}^{\pi/2}\cos\varphi\cos 4\varphi\,d\varphi}{\int\limits_{0}^{\pi/2}\cos^2 4\varphi\,d\varphi}=-\frac{16}{15\pi}.$$

Подставляя найденные значения коэффициентов в формулу (13), находим

$$Q_4(x)=\frac{2}{\pi}+\frac{4}{3\pi}\left(x^2-\frac{1}{2}\right)-\frac{16}{15\pi}\left(x^4-x^2+\frac{1}{8}\right)=\frac{2}{15\pi}(-16x^4+36x^2+3).$$

Полезно заметить следующее. Так как вес $p(x)=\frac{1}{\sqrt{1-x^2}}$ возрастает при приближении к концам отрезка, то отсюда можно заключить, что приближения, получаемые с помощью полиномов Чебышева $T_n(x)$, учитывают в большей степени значения аппроксимируемой функции $f(x)$ у концов отрезка $[-1, 1]$.

Сравним, например, результат последнего примера с примером 1 из § 10, в котором функция $f(x)=|x|$ аппроксимировалась многочленом

$$\tilde{Q}_4(x)=\frac{15}{128}(1+14x^2-7x^4), \qquad |x|\leqslant 1,$$

при весе $p(x)=1$.

В нашем случае при $x=1$ имеем

$$|x|-Q_4(x)=1-\frac{2}{15\pi}(-16+36+3)=0{,}0239,$$

тогда как

$$|x|-\tilde{Q}_4(x)=1-\frac{15}{128}(1+14-7)=0{,}0625.$$

§ 13. Понятие о равномерном приближении функций

Квадратичная аппроксимация функций основывалась на способе наименьших квадратов. Уточняя понятие квадратичного отклонения, введем соответствующее расстояние $\tilde{\Delta}$ между данной непрерывной функцией $f(x)$ и непрерывным аппроксимирующим обобщенным полиномом $Q(x)$ (вообще говоря, более простой природы), так называемое *среднее квадратичное отклонение*.

Определение 1. Под *средним квадратичным отклонением* функций $f(x)$ и $Q(x)$ на множестве точек $X=\{x_1, x_2, \ldots, x_n\}$ понимается число

$$\tilde{\Delta}_n=\sqrt{\frac{1}{n}\sum_{i=1}^{n}[f(x_i)-Q(x_i)]^2}. \tag{1}$$

Если аппроксимация интегральная, то среднее квадратичное отклонение на отрезке $[a, b]$ определяется формулой

$$\tilde{\Delta}=\sqrt{\frac{1}{b-a}\int_a^b[f(x)-Q(x)]^2\,dx}. \tag{2}$$

Формулу (2) можно рассматривать как предельный случай формулы (1) при $n\to\infty$. Действительно, выбирая на отрезке $[a, b]$ систему равноотстоящих точек $a=x_1, x_2, \ldots, x_n=b$, где

$$\Delta x_i=x_{i+1}-x_i=\frac{b-a}{n} \qquad (i=1, 2, \ldots, n-1),$$

будем иметь

$$\tilde{\Delta}_n^2 = \frac{1}{n}\sum_{i=1}^{n}[f(x_i)-Q(x_i)]^2 = \frac{1}{b-a}\sum_{i=1}^{n}[f(x_i)-Q(x_i)]^2\,\Delta x_i.$$

Отсюда при $n \to \infty$ получим

$$\tilde{\Delta}^2 = \lim_{n\to\infty}\tilde{\Delta}_n^2 = \frac{1}{b-a}\int_a^b [f(x)-Q(x)]^2\,dx.$$

Если среднее квадратичное отклонение $\tilde{\Delta}$ мало, то для «подавляющего большинства» значений аргумента x из отрезка $[a, b]$ (т. е.

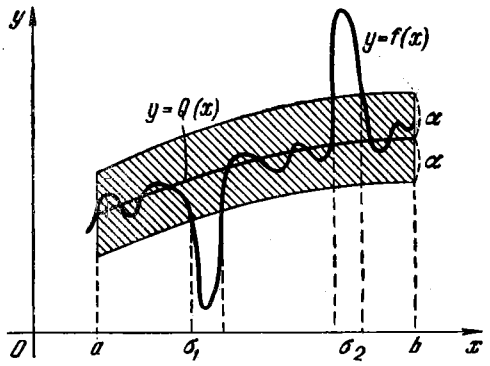

Рис. 8.

в «среднем» на $[a, b]$) абсолютная величина $|f(x)-Q(x)|$ также мала при $a \leqslant x \leqslant b$. Более точно, пусть $|f(x)-Q(x)|$ имеет на $[a, b]$ конечное число экстремумов и α — заданное положительное число. Обозначим через $\sigma = \sigma_1 + \sigma_2 + \ldots + \sigma_k$ максимальную систему непересекающихся отрезков из $[a, b]$ таких, что $|f(x)-Q(x)| \geqslant \alpha$ при $x \in \sigma_i$ $(i = 1, 2, \ldots, k)$ (рис. 8), и пусть ω — сумма длин этих отрезков. Если $\tilde{\Delta} < \varepsilon$, то имеем

$$\varepsilon^2(b-a) > \int_a^b [f(x)-Q(x)]^2\,dx \geqslant \int_\sigma [f(x)-Q(x)]^2\,dx \geqslant \alpha^2\omega.$$

Отсюда $\omega < (b-a)\left(\frac{\varepsilon}{\alpha}\right)^2$, и, следовательно, ω — сколь угодно малое число, если ε достаточно мало.

Таким образом, если $\tilde{\Delta} < \varepsilon$, где ε достаточно мало, то на отрезке $[a, b]$, за исключением, быть может, множества точек σ, сколь угодно малой линейной меры ω, выполнено неравенство

$$|f(x) - Q(x)| < \alpha,$$

где α — произвольное, заранее заданное положительное число.

Во многих случаях, например при обработке результатов наблюдений, квадратичное приближение является приемлемым, так как оно

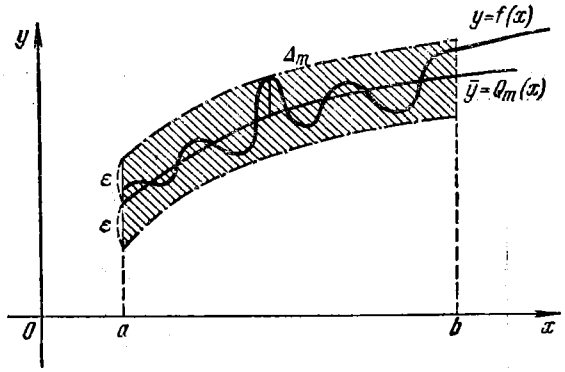

Рис. 9.

сглаживает отдельные локальные неправильности функции $f(x)$ (возникшие, возможно, от ошибок наблюдений) и дает достаточно точное общее представление о протекании соответствующего процесса. Однако иногда для приближения ставят более жесткие условия, а именно требуется гарантировать, чтобы на всем отрезке $[a, b]$ отклонение функций $f(x)$ и $Q(x)$ было меньше заданной величины. Поэтому введем другое расстояние между функциями, так называемое абсолютное отклонение.

Определение 2. *Абсолютным отклонением* на $[a, b]$ обобщенного полинома $Q_m(x)$ от данной непрерывной функции $f(x)$ называется число

$$\Delta(f, Q_m) \equiv \Delta_m = \max_{a \leqslant x \leqslant b} |f(x) - Q_m(x)|. \qquad (3)$$

Если $\Delta_m \leqslant \varepsilon$, то из формулы (3) следует $|f(x) - Q_m(x)| \leqslant \varepsilon$ для всех точек x на отрезке $[a, b]$ (рис. 9).

В этом случае говорят, что обобщенный полином $Q_m(x)$ на отрезке $[a, b]$ *равномерно приближает функцию* $f(x)$ с точностью до ε.

Для случая обычных полиномов

$$Q_m(x) = \sum_{k=0}^{m} a_k x^k$$

справедлива важная аппроксимационная теорема Вейерштрасса, которую мы приводим без доказательства (см. [1]).

Теорема Вейерштрасса. *Если функция $f(x)$ непрерывна на отрезке $[a, b]$, то, как бы мало ни было положительное число ε, найдется полином $Q_m(x)$ достаточно высокой степени m, абсолютное отклонение которого от данной функции $f(x)$ на отрезке $[a, b]$ меньше, чем ε, т. е. для всех точек $x \in [a, b]$ имеет место неравенство*

$$|f(x) - Q_m(x)| < \varepsilon.$$

В частном случае, если функция $f(x)$ аналитическая на отрезке $[a, b]$, т. е. разлагается на этом отрезке в равномерно сходящийся степенной ряд (ряд Тейлора)

$$f(x) = \sum_{k=0}^{\infty} a_k (x-c)^k,$$

где

$$c \in [a, b], \qquad a_k = \frac{1}{k!} f^{(k)}(c) \qquad (k = 0, 1, 2, \ldots),$$

то за полином $Q_m(x)$ можно взять отрезок ряда Тейлора

$$Q_m(x) = \sum_{k=0}^{m} a_k (x-c)^k,$$

причем степень m подбирается в зависимости от заданной точности ε.

Пусть степень m полинома $Q_m(x)$ фиксирована и задача состоит в том, чтобы приблизить данную непрерывную функцию $f(x)$ полиномом фиксированной степени m *наилучшим образом* на заданном множестве X. Это значит, что коэффициенты $a_0, a_1, \ldots, a_m$ полинома

$$Q_m(x) = a_0 + a_1 x + \ldots + a_m x^m \tag{4}$$

следует подобрать так, чтобы величина

$$\Delta_m = \max_{x \in X} |f(x) - Q_m(x)| \tag{5}$$

была минимальной.

Такой полином $Q_m(x)$, дающий минимум величине Δ_m, называется *полиномом наилучшего равномерного приближения* или *полиномом, наименее отклоняющимся от функции $f(x)$ на множестве X.*

При этом минимальное отклонение

$$E_m[f, X] = \min \Delta_m \tag{6}$$

называется *наименьшим отклонением* функции $f(x)$ на множестве X.

В теории приближения функций доказывается существование и единственность полинома наилучшего приближения для любой непрерывной функции [5]—[8].

Теорема 1. *Для любой функции $f(x)$, непрерывной на замкнутом ограниченном множестве X *), и любого натурального числа m существует полином $Q_m(x)$ степени не выше m, обладающий минимальным отклонением*

$$\Delta(f, Q_m) = E_m[f, X], \tag{7}$$

причем такой полином единственный.

В некоторых случаях на полином $Q_m(x)$ накладывают дополнительные ограничения, например, полагают старший коэффициент $a_m = 1$. Тогда

$$Q_m(x) = a_0 + a_1 x + a_2 x^2 + \ldots + x^m. \tag{8}$$

В частности, пусть $f(x) \equiv 0$; тогда полином $Q_m(x)$, дающий минимум величине

$$\rho_m = \max_{a \leqslant x \leqslant b} |Q_m(x)|,$$

называется *полиномом, наименее отклоняющимся от нуля* **) на данном отрезке $[a, b]$.

Покажем, что этим свойством обладает полином Чебышева.

Теорема 2. *Полином Чебышева степени m ($m > 1$) наименее отклоняется от нуля на отрезке $[-1, 1]$ по сравнению с другим полиномом степени m и со старшим коэффициентом, равным единице.*

Доказательство. Допустим, что существует полином $Q_m(x)$ со старшим коэффициентом, равным единице, и такой, что

$$\max_{-1 \leqslant x \leqslant 1} |Q_m(x)| \leqslant \max_{-1 \leqslant x \leqslant 1} |T_m(x)| = \rho_m. \tag{9}$$

В силу теоремы 3 § 12 и ее следствия $\rho_m = 1/2^{m-1}$, причем на отрезке $[-1, 1]$ существует $m+1$ точек $1 = \xi_0, \xi_1, \ldots, \xi_m = -1$ таких, что (рис. 10)

$$T_m(\xi_k) = (-1)^k \rho_m. \tag{10}$$

Рассмотрим разность $R(x) = T_m(x) - Q_m(x)$. Так как старшие коэффициенты полиномов $T_m(x)$ и $Q_m(x)$ одинаковы, то $R(x)$ есть полином степени не выше $m-1$.

Из условия (9) и формулы (10) вытекает, что

$$R(\xi_0) \geqslant 0, \; R(\xi_1) \leqslant 0, \; \ldots, \; R(\xi_m)(-1)^m \geqslant 0. \tag{11}$$

На основании неравенств (11) заключаем, что полином $R(x)$ имеет по меньшей мере один нуль на каждом из отрезков $[\xi_1, \xi_0]$,

---

*) В дальнейшем X будет представлять собой или отрезок, или конечную систему точек $x_1, x_2, \ldots, x_n$.

**) Задача о разыскании полинома вида (4), наименее отклоняющегося от нуля, имеет тривиальное решение $Q_m(x) \equiv 0$.

$[\xi_2, \xi_1], \ldots, [\xi_m, \xi_{m-1}]$ (на рис. 10 эти нули суть абсциссы точек пересечения кривых $y = T_m(x)$ и $y = Q_m(x)$), причем общее количество нулей с учетом их кратностей будет не меньше m.

Так как степень полинома $R(x)$ не превышает $m-1$, то это возможно лишь в случае, если $R(x) \equiv 0$, т. е. $Q_m(x) \equiv T_m(x)$. Теорема доказана.

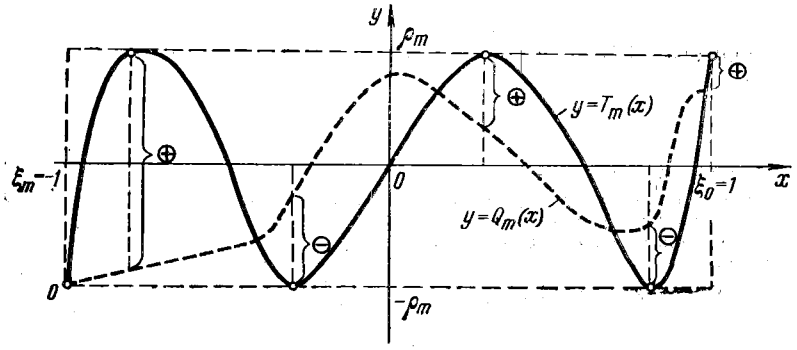

Рис. 10.

Следствие 1. Для любого полинома $Q_m(x)$ степени m со старшим коэффициентом, равным единице, имеет место неравенство

$$\max |Q_m(x)| \geqslant \frac{1}{2^{m-1}} \quad (-1 \leqslant x \leqslant 1).$$

Следствие 2. Для функции $f(x) = x^m$ $(m \geqslant 1)$ полиномом наилучшего приближения $Q_{m-1}(x) = a_0 + a_1 x + \ldots + a_{m-1} x^{m-1}$ степени $m-1$ на отрезке $[-1, 1]$ является

$$Q_{m-1}(x) = x^m - T_m(x), \tag{12}$$

где $T_m(x)$ — полином Чебышева.

Действительно, согласно смыслу задачи разность

$$f(x) - Q_{m-1}(x) = x^m - a_{m-1}x^{m-1} - \ldots - a_1 x - a_0$$

есть полином, наименее отклоняющийся от нуля на отрезке $[-1, 1]$, т. е. представляет собой полином Чебышева $T_m(x)$. Отсюда непосредственно вытекает формула (12).

Замечание. Пользуясь теоремой 2, легко построить наименее отклоняющийся от нуля на данном отрезке $[a, b]$ полином $\hat{T}_m(x)$ степени m со старшим коэффициентом, равным единице [7].

Действительно, подстановка

$$x = \frac{b+a}{2} + \frac{b-a}{2} t$$

преобразует отрезок $a \leqslant x \leqslant b$ в отрезок $-1 \leqslant t \leqslant 1$, причем старший коэффициент (при t^m) будет равен $\left(\frac{b-a}{2}\right)^m$. Отсюда

$$\hat{T}_m(x) = \left(\frac{b-a}{2}\right)^m T_m(t) = \left(\frac{b-a}{2}\right)^m T_m\left(\frac{x-\frac{b+a}{2}}{\frac{b-a}{2}}\right). \qquad (13)$$

Из формулы (13) следует, что отклонение полинома $\hat{T}_m(x)$ от нуля равно

$$\hat{E}_m = \left(\frac{b-a}{2}\right)^m \frac{1}{2^{m-1}} = 2\left(\frac{b-a}{4}\right)^m.$$

Пример 1. С помощью полинома первой степени $Q_1(x) = Ax + B$ наилучшим образом равномерно приблизить функцию $f(x) = x^2$ на отрезке $[0, 1]$.

Решение. Требуется определить коэффициенты A и B так, чтобы величина

$$E_1 = \max_{0 \leqslant x \leqslant 1} |x^2 - Ax - B| \qquad (14)$$

была наименьшей. Следовательно, полином $Q_2(x) = x^2 - Ax - B$ наименее отклоняется от нуля на отрезке $[0, 1]$. Полагая $a = 0$ и $b = 1$, в силу формулы (13) получаем

$$Q_2(x) = \left(\frac{1}{2}\right)^2 T_2\left(\frac{x - {}^1/_2}{{}^1/_2}\right) = \frac{1}{4} T_2(2x-1) = \frac{1}{4}\left[(2x-1)^2 - \frac{1}{2}\right] = x^2 - x + \frac{1}{8}$$

(см. § 12). Отсюда

$$Q_1(x) = x - {}^1/_8,$$

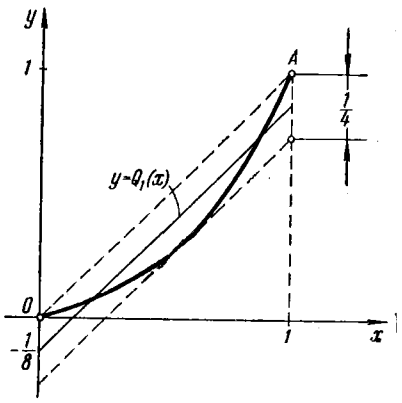

Рис. 11.

причем $E_1 = 2({}^1/_4)^2 = {}^1/_8$. Заметим, что отклонение E_1 реализуется в трех точках:

$$Q_1(0) = Q_1({}^1/_2) = Q_1(1) = E_1.$$

Это явление характерно для полиномов наилучшего приближения.

Геометрически график функции $y = Q_1(x)$ представляет собой среднюю параллель между секущей, проходящей через крайние точки $O(0, 0)$ и $A(1, 1)$, и касательной, параллельной этой секущей (см. рис. 11, где кривая — график функции $y = x^2$).

Теорема Чебышева об альтернансе. *Среди полиномов степени не выше m полином $Q_m(x)$ является полиномом наилучшего равномерного приближения непрерывной функции $f(x)$ на данном множестве X тогда и только тогда, когда на X существует система $(m+2)$-х точек*

$$\xi_0 < \xi_1 < \ldots < \xi_{m+1}$$

(чебышевский альтернанс) таких, что разность $f(x)-Q_m(x)$ *принимает поочередно значения* L *и* $-L$, *где*

$$L=\max_X |f(x)-Q_m(x)|. \tag{15}$$

Пример 2. Для функции $f(x)=\cos 2x$ $(0\leqslant x\leqslant 2\pi)$ полиномом третьей степени (в широком смысле) наилучшего равномерного приближения является $Q_3(x)\equiv 0$.

Действительно, здесь $L=\max\limits_{x\in[0,\ 2\pi]}|\cos 2x-0|=1$ и разность $f(x)-Q_3(x)=$ $=\cos 2x$ принимает последовательно значения $+1$ и -1 в пяти точках $0,\ \pi/2,\ \pi,\ 3\pi/2,\ 2\pi$.

Задача построения точного альтернанса весьма трудна. Приближенное построение альтернанса и связанное с ним построение полинома наилучшего приближения предложено Ремезом [7], [5].

ЛИТЕРАТУРА К ГЛАВЕ I

[1] Гончаров В. Л., Теория приближения и интерполирования функций, Гостехиздат, 1954, гл. I—III.
[2] Демидович Б. П., Марон И. А., Основы вычислительной математики, «Наука», 1966, гл. VIII, IX, XIV.
[3] Милн В. Э., Численный анализ, ИЛ, 1951, гл. IX, X.
[4] Янке Е. и Эмде Ф., Таблицы функций с формулами и кривыми, изд. 3, Физматгиз, 1959, ч. II, гл. VII.
[5] Березин И. С., Жидков Н. П., Методы вычислений, «Наука», 1966, т. 1, гл. IV.
[6] Натансон И. П., Конструктивная теория функций, Гостехиздат, 1949, гл. II.
[7] Ремез Е. Я., Общие вычислительные методы чебышевского приближения, Изд. АН УССР, 1957, гл. I.
[8] Ахиезер Н. И., Лекции по теории аппроксимации, «Наука», 1965.

ГЛАВА II

ЭМПИРИЧЕСКИЕ ФОРМУЛЫ

§ 1. Вводные замечания

Пусть, изучая функциональную зависимость

$$y = f(x), \tag{1}$$

мы произвели ряд измерений величин x и y и в результате получили таблицу значений

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

или график, связывающий значения x и y.

Если аналитическое выражение функции $f(x)$ неизвестно или весьма сложно, то возникает практически важная задача: найти *эмпирическую формулу*

$$\tilde{y} = \tilde{f}(x), \tag{2}$$

значения которой при $x = x_i$ возможно мало отличались бы от опытных данных y_i $(i = 1, 2, \ldots, n)$. В такой постановке наша задача весьма неопределенна; поэтому обычно по ряду соображений указывают достаточно узкий класс функций K (например, множество функций линейных, степенных, показательных и т. п.), которому должна принадлежать искомая функция $\tilde{f}(x)$, и дело, таким образом, сводится к нахождению лишь наилучших значений параметров. Во многих случаях класс K определяется требованием простоты эмпирической формулы; иногда этот класс подсказывается самой природой явления.

Геометрически задача построения эмпирической формулы состоит в проведении кривой Г вида (2) из некоторого класса K, «возможно ближе» примыкающей к системе точек $M_i(x_i, y_i)$ $(i = 1, 2, \ldots, n)$

(рис. 12). Разумеется, при этом должен быть выяснен точный математический смысл понятия «близости» кривой Г к конфигурации точек M.

Заметим, что задача построения эмпирической формулы отлична от задачи интерполирования (гл. I, § 2). При интерполировании отыскивается функция из данного класса функций (например, полиномов заданной степени), значения которой в заданных точках x_i совпадали бы с табличными значениями y_i ($i = 1, 2, \ldots, n$). При нахождении эмпирической формулы не требуется, чтобы значения $\tilde{f}(x_i)$ совпадали с y_i, достаточно, чтобы разность $\tilde{f}(x_i) - f(x_i)$ была мала в известном смысле в данной области. Следует иметь в виду также, что сами исходные эмпирические данные x_i и y_i, как правило, являются приближенными и содержат ошибки. Поэтому интерполяционная формула, повторяющая эти ошибки, не говоря даже о ее сложности, не является идеальным решением поставленной задачи; часто простая эмпирическая формула, сглаживающая местные неправильности, может лучше отобразить действительность.

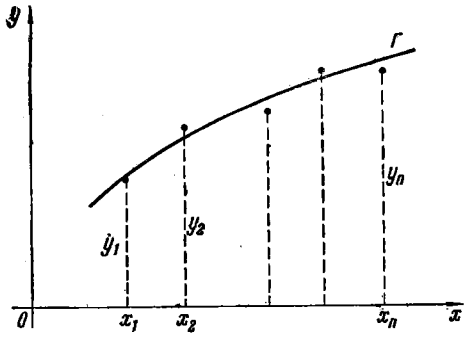

Рис. 12.

Построение эмпирической формулы слагается из двух этапов: 1) выяснение общего вида этой формулы и 2) определение наилучших параметров ее.

Если неизвестен характер зависимости между данными величинами x и y, то вид эмпирической формулы является произвольным. Предпочтение отдается простым формулам, обладающим хорошей точностью. Если отсутствуют сведения о промежуточных данных, то обычно предполагается, что эмпирическая функция аналитическая без точек разрыва и график ее — плавная кривая. Нельзя указать общего метода для нахождения наилучшего типа формулы, соответствующей опытным данным.

Удачный подбор эмпирической формулы в значительной мере зависит от опыта и искусства составителя.

В некоторых случаях выбор типа эмпирической формулы может быть произведен на основе теоретических представлений о характере изучаемой зависимости. В других случаях удается подобрать такую формулу, сравнивая кривую, построенную по данным наблюдения в декартовых координатах или в специальных системах координат

(полулогарифмической, логарифмической и т. д.), с образцами известных кривых (отдельные неправильности при этом игнорируются). Для облегчения выбора полезно использовать специальные альбомы кривых. При известном навыке по положению точек, определяющих некоторую гладкую кривую, можно примерно угадать общий вид зависимости.

Во многих случаях можно ограничиться полиномом

$$y = \sum_{k=0}^{m} a_k x^k. \tag{3}$$

Нередко употребляются другие элементарные функции (дробно-линейная, степенная, показательная, логарифмическая и т. п.). В дальнейшем будут указаны приемы, облегчающие выбор вида эмпирической формулы.

Что касается определения наилучших значений параметров, входящих в эмпирическую формулу, то эта задача более легкая и решается регулярными методами.

Эмпирические формулы не претендуют на роль законов природы, а являются лишь гипотезами, более или менее удовлетворительно согласующимися с наблюденными опытными данными. Однако значение их весьма велико; в истории науки известны многочисленные примеры того, как получение удачной эмпирической формулы приводило к большим научным открытиям.

Замечание. При построении эмпирической формулы можно предполагать, что исходные данные (x_i, y_i) $(i = 1, 2, \ldots, n)$ положительны.

Действительно, если бы, например, все $x_i < 0$ (или все $y_i < 0$), то достаточно рассмотреть таблицу значений $(-x_i, y_i)$ (или соответственно $(x_i, -y_i)$). Аналогично при $x_i < 0$ и $y_i < 0$ достаточно построить эмпирическую формулу для таблицы $(-x_i, -y_i)$.

Пусть теперь имеем общий случай, когда знаки значений x_i и y_i переменные. Так как таблица значений (x_i, y_i) конечна, то всегда можно подобрать положительные числа m и n такие, что

$$\xi_i = m + x_i > 0, \qquad \eta_i = n + y_i > 0.$$

Отсюда получаем, что решение поставленной задачи сводится к нахождению эмпирической формулы для системы положительных значений (ξ_i, η_i).

Поэтому в дальнейшем мы, как правило, будем рассматривать таблицы с положительными элементами.

Заметим, что мы в дальнейшем не будем касаться вопросов, связанных с оценкой доброкачественности исходных данных, в частности, не будем учитывать ошибки этих данных.

§ 2. Линейная зависимость

Пусть для переменных x и y известны их значения x_i и y_i ($i=1, 2, \ldots, n$), расположенные в порядке возрастания первой переменной: $x_1 < x_2 < \ldots < x_n$.

На координатной плоскости Oxy построим систему точек $M_i(x_i, y_i)$ (для удобства масштабы на осях Ox и Oy могут быть выбраны разными). Если окажется, что эти точки располагаются примерно на некоторой прямой линии L, то естественно предположить, что зависимость между x и y *линейная*:

$$y = ax + b, \tag{1}$$

где a и b — постоянные.

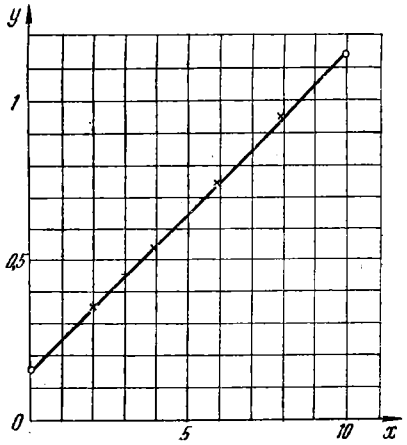

Рис. 13.

Прикладывая прозрачную линейку (или натягивая нить) так, чтобы положение ее было возможно близким к каждой из точек, опытным путем можно найти наилучшее положение прямой L. Построение прямой L обычно производится на бумаге с миллиметровой сеткой, причем результат не является однозначно определенным.

Для нахождения параметров a и b измеряют координаты двух точек (не обязательно из данных) $N_1(\tilde{x}_1, \tilde{y}_1)$ и $N_2(\tilde{x}_2, \tilde{y}_2)$, лежащих на прямой L (например, находят точки пересечения ее с осями координат) и по возможности далеко удаленных друг от друга. Тогда a и b могут быть определены из системы уравнений *(способ выбранных точек)*

$$\tilde{y}_1 = a\tilde{x}_1 + b, \qquad \tilde{y}_2 = a\tilde{x}_2 + b. \tag{2}$$

Результат исключения коэффициентов a и b из уравнений (2) может быть записан в виде определителя:

$$\begin{vmatrix} y & x & 1 \\ \tilde{y}_1 & \tilde{x}_1 & 1 \\ \tilde{y}_2 & \tilde{x}_2 & 1 \end{vmatrix} = 0. \tag{3}$$

Геометрически уравнение (3) представляет собой уравнение прямой, проходящей через точки N_1 и N_2.

Метод выбранных точек весьма нагляден, но обладает малой точностью. В дальнейшем (см. §§ 5—8) будут указаны более точные аналитические методы определения коэффициентов a и b.

Пример 1. Подобрать эмпирическую формулу для табличных данных:

x	2	4	6	8
y	0,350	0,573	0,725	0,947

Решение. Построив соответствующие точки $M_i(x_i, y_i)$ $(i=1, 2, 3, 4)$ в подходящем масштабе, убеждаемся, что они располагаются примерно на прямой линии $y=ax+b$ (рис. 13).

Из графика находим отрезок, отсекаемый прямой на оси ординат: $b=0,147$. Кроме того, при $x=10$ имеем $y=1,138$. Поэтому

$$a=\frac{1,138-0,147}{10}=0,099.$$

Можно принять

$$\tilde{y}=0,099x+0,147. \tag{4}$$

Для сравнения в таблице 10 приведены значения, полученные по формуле (4), и разности $\Delta=y-\tilde{y}$.

Таблица 10

Значения эмпирической функции $\tilde{y}$

x	2	4	6	8
$\tilde{y}$	0,345	0,543	0,741	0,939
$\Delta=y-\tilde{y}$	0,005	0,050	−0,016	0,008

Легко дать аналитический критерий для прямолинейности ряда точек $M_i(x_i, y_i)$.

Положим

$$\Delta x_i=x_{i+1}-x_i, \qquad \Delta y_i=y_{i+1}-y_i \tag{5}$$

и

$$k_i=\frac{\Delta y_i}{\Delta x_i} \quad (i=1, 2, \ldots, n-1). \tag{6}$$

Если $k_i=\text{const}$, то точки $M_i(x_i, y_i)$, очевидно, лежат *точно* на одной прямой линии. Если

$$k_1\approx k_2\approx\ldots\approx k_{n-1}, \tag{7}$$

то точки $M_i(x_i, y_i)$ приблизительно расположены на прямой. В зависимости от точности выполнения соотношений (7) решается вопрос: следует или не следует искать эмпирическую зависимость между величинами x и y в виде линейной функции?

В частности, если значения x_i равноотстоящие, т. е. $\Delta x_i = \text{const}$, то достаточно убедиться, что значения Δy_i являются также постоянными (или почти постоянными).

Пример 2. Проверить на прямолинейность следующий ряд точек:

i	1	2	3	4	5	6
x_i	0,5	1,0	1,5	2,0	2,5	3,0
y_i	0,62	1,64	2,58	3,70	5,02	6,04

Решение. Так как значения x_i равноотстоящие, то составляем таблицу разностей Δy_i (таблица 11).

Таблица 11

Проверка на прямолинейность

i	1	2	3	4	5
Δy_i	1,02	0,94	1,12	1,32	1,02

Мы видим, что значения Δy_i близки друг к другу (за исключением Δy_4). Поэтому в качестве грубой эмпирической формулы можно выбрать линейную.

§ 3. Метод выравнивания

Пусть для переменных x и y их соответствующие значения x_i и y_i $(i = 1, 2, \ldots, n)$ таковы, что точки $M_i(x_i, y_i)$ не располагаются на прямой линии. Тогда во многих случаях, вводя новые переменные

$$X = \varphi(x, y), \quad Y = \psi(x, y), \tag{1}$$

можно добиться того, чтобы преобразованные точки $N_i(X_i, Y_i)$, где $X_i = \varphi(x_i, y_i)$ и $Y_i = \psi(x_i, y_i)$, лежали на некоторой прямой линии плоскости OXY *(метод выравнивания)*. Обязательным требованием при этом является взаимная однозначность преобразования (1).

Рассмотрим, например, нелинейную зависимость вида

$$\psi(y) = a\varphi(x) + b, \tag{2}$$

где a и b — постоянные, $\varphi(x)$ и $\psi(y)$ — строго монотонные функции. На плоскости Oxy функция (2) изображается некоторой кривой.

Вводя новые переменные $X=\varphi(x)$ и $Y=\psi(y)$, будем иметь

$$Y=aX+b, \tag{3}$$

и, следовательно, при наличии зависимости (2) точки $N_i(\varphi(x_i), \psi(y_i))$ $(i=1, 2, \ldots, n)$ на новой координатной плоскости OXY располагаются на прямой линии. Обратно, если при построении на плоскости OXY обнаружится, что точки N_i практически лежат на прямой линии, то между переменными x и y имеет место зависимость (2).

Функции $\varphi(x)$ и $\psi(y)$ обычно находятся методом проб, на основании обнаружения сходства линии (2) с известными кривыми, или проверкой выполнения соответствующих аналитических критериев (см. § 4).

К виду (2) сводится, например, *степенная зависимость*

$$y=cx^a, \tag{4}$$

где a и c — постоянные, причем $x>0$ и $y>0$.

Логарифмируя формулу (4), будем иметь

$$\lg y=a\lg x+\lg c.$$

Отсюда, полагая $X=\lg x$, $Y=\lg y$, получим

$$Y=aX+b, \qquad b=\lg c. \tag{5}$$

Таким образом, степенная зависимость (4) между переменными x и y обнаруживается, если точки $N_i(\lg x_i, \lg y_i)$ $(i=1, 2, \ldots, n)$ лежат на одной прямой в плоскости OXY. Для построения прямой (5) выгодно использовать логарифмическую бумагу, представляющую собой неравномерную сетку, построенную на горизонтальной и вертикальной логарифмических шкалах. Поэтому, откладывая на осях численные значения x_i и y_i, мы получаем точку N_i с координатами $\lg x_i$ и $\lg y_i$ (в известном масштабе), т. е. избавляемся от необходимости вычислять значения логарифмов исходных координат.

Заметим, что в логарифмической координатной сетке OXY началом координат служит точка $x=1$, $y=1$. Поэтому исходные данные x_i и y_i (которые должны быть положительными) выгодно преобразовать так, чтобы имели место соотношения $x_i\geqslant 1$ и $y_i\geqslant 1$ (например, умножить их на подходящий положительный множитель). Полагая $x=1$ в соотношении (4), будем иметь $y=c$; таким образом, постоянная c представляет собой ординату точки пересечения с осью OY соответствующего прямолинейного графика в плоскости OXY.

Пример 1. Табличные данные

x	10	20	30	40	50	60	70	80
y	1,06	1,33	1,52	1,68	1,81	1,91	2,01	2,11

отвечают формуле вида

$$y = cx^a. \tag{6}$$

Найти параметры a и c.

Решение. Нанося данные точки на логарифмическую бумагу, замечаем, что эти точки с достаточной точностью располагаются на прямой

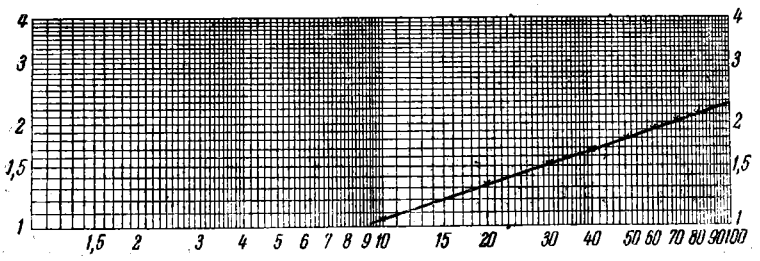

Рис. 14.

линии (рис. 14). Тем самым подтверждается наше предположение (6) о характере зависимости между x и y.

Для определения параметров a и c воспользуемся крайними точками A (10; 1,06) и B (80; 2,11).

Логарифмируя соотношение (6), получим

$$\lg y = \lg c + a \lg x.$$

Отсюда, подставляя координаты точек A и B, находим

$$\lg 1{,}06 = \lg c + a \lg 10, \qquad \lg 2{,}11 = \lg c + a \lg 80$$

или

$$\lg c + a = 0{,}0253, \qquad \lg c + 1{,}9031\, a = 0{,}3243;$$

следовательно,

$$a = \frac{0{,}3243 - 0{,}0253}{1{,}9031 - 1} = \frac{0{,}2990}{0{,}9031} = 0{,}3311,$$

$$\lg c = 0{,}0253 - 0{,}3311 = -0{,}3058 = \bar{1}{,}6942; \quad c = 0{,}4945.$$

Таким образом,

$$\tilde{y} = 0{,}4945 x^{0{,}3311}. \tag{7}$$

В таблице 12 приведены значения $\tilde{y}=\tilde{y}(x)$, вычисленные по формуле (7), и их отклонения $\Delta=y-\tilde{y}$ от табличных данных.

Таблица 12

Значения эмпирической функции (7)

x	10	20	30	40	50	60	70	80
$\tilde{y}$	1,060	1,320	1,525	1,677	1,806	1,919	2,012	2,110
$\Delta=y-\tilde{y}$	0	0,010	— 0,005	0,003	0,004	— 0,009	— 0,002	0

Другой важный случай представляет *показательная зависимость*

$$y=ce^{ax}, \quad c>0. \tag{8}$$

Логарифмируя, будем иметь

$$\ln y=ax+\ln c.$$

Отсюда, полагая $X=x$, $Y=\ln y$, получим линейную зависимость

$$Y=aX+b, \quad b=\ln c.$$

Если имеется подозрение на показательную зависимость (8), то для построения в плоскости OXY удобно использовать полулогарифмическую бумагу с равномерной шкалой по оси OX и логарифмической шкалой по оси OY.

Начало координат в полулогарифмической сетке соответствует точке $x=0$, $y=1$; поэтому удобно брать $y_i \geqslant 1$. Для прямолинейного графика в плоскости OXY, соответствующего формуле (8), постоянная c представляет собой ординату точки пересечения этого графика с осью OY ($x=0$, $y=c$).

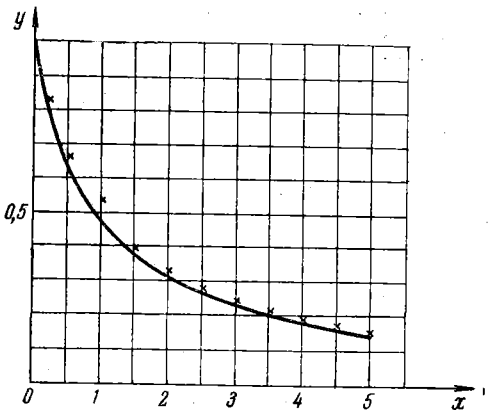

Рис. 15.

В общем случае выбор вида эмпирической формулы облегчается знакомством с графиками элементарных функций.

Пример 2. Результаты измерений выражаются таблицей [9]:

x	0	0,2	0,5	1	1,5	2	2,5	3	3,5	4	4,5	5
y	1	0,833	0,667	0,540	0,400	0,333	0,286	0,250	0,222	0,200	0,182	0,167

Составить подходящую эмпирическую формулу.

Решение. Нанося точки $M_i(x_i, y_i)$ на координатную сетку Oxy, обнаруживаем, что они располагаются примерно на кривой, похожей на ветвь гиперболы, асимптотически приближающуюся к оси Ox (рис. 15). Поэтому полагаем

$$y=\frac{1}{ax+b}. \tag{9}$$

Отсюда

$$\frac{1}{y}=ax+b,$$

т. е. имеем зависимость вида (2). Введя новые переменные $X=x$, $Y=\frac{1}{y}$, получим линейную зависимость

$$Y=aX+b. \tag{10}$$

Соответствующие данные приведены в таблице 13.

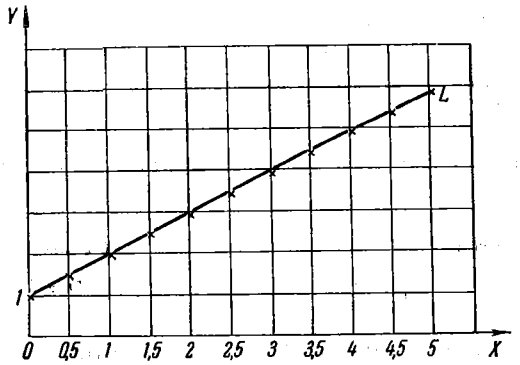

Рис. 16.

Таблица 13

Подбор эмпирической формулы методом выравнивания

X	0	0,2	0,5	1	1,5	2	2,5	3	3,5	4	4,5	5
Y	1	1,2	1,5	2,0	2,5	3,0	3,5	4,0	4,5	5,0	5,5	6

Построив в плоскости OXY точки $N_i(X_i, Y_i)$, убеждаемся, что они лежат приблизительно на прямой линии L (рис. 16).

Следовательно, наша гипотеза о характере зависимости оправдалась и вид формулы (9) выбран правильно. По рис. 16 определяем параметры прямой L:

$$b=1, \quad a=\frac{6-1}{5-0}=1.$$

Таким образом, искомая эмпирическая формула есть $y=\frac{1}{x+1}$.

§ 4. Квадратичная (параболическая) зависимость

Если для данных значений (x_i, y_i) $(i=1, 2, \ldots, n)$ не оправдывается линейная зависимость, то можно испробовать более общую *квадратичную зависимость*

$$y=ax^2+bx+c \quad (a \neq 0). \qquad (1)$$

Формула (1) имеет место, когда точки $M_i(x_i, y_i)$ располагаются на отрезке параболы с вертикальной осью (рис. 17). Для обнаружения этого факта с надлежащей тщательностью проводим плавную кривую Г, вблизи которой группируются данные точки $M_i(x_i, y_i)$, и выбираем на ней точку $N(x_0, y_0)$, по возможности совпадающую с одной из точек $M_k(x_k, y_k)$ $(1 \leqslant k \leqslant n)$.

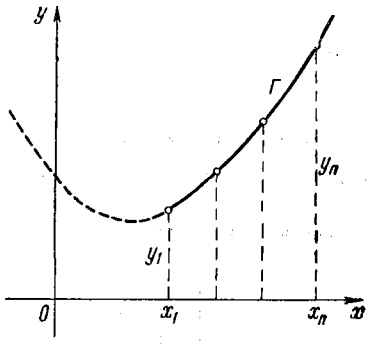

Рис. 17.

Предполагая, что кривая Г есть парабола (1), будем иметь

$$y_0=ax_0^2+bx_0+c. \qquad (2)$$

Вычитая из уравнения (1) равенство (2), находим

$$y-y_0=a(x^2-x_0^2)+b(x-x_0)$$

или

$$y-y_0=a(x-x_0)^2+b_1(x-x_0), \qquad (3)$$

где $b_1=b+2ax_0$.

Если теперь ввести новые переменные

$$X=x-x_0, \quad Y=\frac{y-y_0}{x-x_0}, \qquad (4)$$

то из уравнения (3) получим линейную зависимость

$$Y=aX+b_1. \qquad (5)$$

Таким образом, параболе (1) на плоскости Oxy соответствует прямая (5) на плоскости OXY (см. § 2). Обратно, если будет

обнаружено (см. § 2), что точки $N_i(X_i, Y_i)$ $(i=1, 2, \ldots, n)$, где

$$X_i = x_i - x_0, \quad Y_i = \frac{y_i - y_0}{x_i - x_0},$$

лежат на прямой (5), то точки $M_i(x_i, y_i)$ расположены на параболе (1), причем

$$b = b_1 - 2ax_0, \quad c = y_0 + ax_0^2 - b_1 x_0. \tag{6}$$

Выведем аналитический критерий для квадратичной зависимости (1). Пусть $M_i(x_i, y_i)$ — данная таблица значений, причем

$$\Delta x_i = x_{i+1} - x_i > 0 \quad (i = 1, 2, \ldots, n-1).$$

При наличии зависимости (1) последовательность $y_1, y_2, \ldots, y_n$ или монотонная, т. е. $\Delta y_i = y_{i+1} - y_i$ $(i = 1, 2, \ldots, n-1)$ сохраняет постоянный знак, или же эта последовательность имеет единственный экстремум, т. е. разность Δy_i лишь один раз меняет знак.

Введем разделенные разности первого и второго порядков (см., например, [15]):

$$[x_i, x_{i+1}] = \frac{\Delta y_i}{\Delta x_i}$$

и

$$[x_i, x_{i+1}, x_{i+2}] = \frac{[x_{i+1}, x_{i+2}] - [x_i, x_{i+1}]}{x_{i+2} - x_i}$$

или

$$[x_i, x_{i+1}, x_{i+2}] = \frac{\Delta\left(\frac{\Delta y_i}{\Delta x_i}\right)}{\Delta_1 x_i}, \tag{7}$$

где $\Delta_1 x_i = x_{i+2} - x_i = \Delta x_i + \Delta x_{i+1}$.

Доказывается [15], что точки $M_i(x_i, y_i)$ расположены на параболе (1) тогда и только тогда, когда сохраняют постоянное значение все разделенные разности второго порядка.

В частности, если значения $x_1, x_2, \ldots, x_n$ равноотстоящие, т. е. $\Delta x_i = h = \text{const}$, то для существования эмпирической квадратичной зависимости (1) необходимо и достаточно, чтобы была постоянной вторая разность $\Delta^2 y_i = y_{i+2} - 2y_{i+1} + y_i \equiv c$ $(i = 1, 2, \ldots, n-2)$, причем тогда $\Delta^2 y_i = 2h^2 a$.

Заметим, что вторые разности весьма чувствительны к отклонениям от параболического закона.

Пример 1. Данную систему значений

x	0	0,5	1	2	4	5	6
y	0	−1,76	−3,00	−3,96	0,24	5,40	13,34

исследовать на квадратичную зависимость.

Решение. Составляем по формуле (7) таблицу разделенных разностей (таблица 14).

Так как вторая разделенная разность примерно постоянна, то можно считать, что между переменными x и y имеет место приближенная квадратичная зависимость.

Таблица 14

Проверка на квадратичную зависимость

x	Δx	y	Δy	$\frac{\Delta y}{\Delta x}$	$\Delta\left(\frac{\Delta y}{\Delta x}\right)$	$\Delta_t x$	$\frac{\Delta\left(\frac{\Delta y}{\Delta x}\right)}{\Delta_t x}$
0		0					
	0,5		— 1,76	— 3,52			
0,5		—1,76			1,04	1	1,04
	0,5		— 1,24	— 2,48			
1		— 3,00			1,52	1,5	1,01
	1		— 0,96	— 0,96			
2		— 3,96			2,92	3	0,97
	2		3,72	1,86			
4		— 0,24			3,78	3	1,26
	1		5,64	5,64			
5		5,40			2,30	2	1,15
	1		7,94	7,94			
6		13,34					

§ 5. Определение параметров эмпирической формулы

Если вид эмпирической формулы выбран, то возникает задача определения наилучших коэффициентов *(параметров)*, входящих в эту формулу.

В общем виде эта задача ставится следующим образом: пусть данная система значений $M_i(x_i, y_i)$ $(i=1, 2, \ldots, n)$ приближенно описывается формулой вида

$$y=\tilde{f}(x;\ a_1,\ a_2,\ \ldots,\ a_m), \tag{1}$$

где $\tilde{f}$ — известная функция и $a_1, a_2, \ldots, a_m$ — неизвестные постоянные, число которых m обычно меньше числа точек M_i, т. е. $m<n$. Требуется определить эти постоянные.

Если значения (x_i, y_i) точно связаны зависимостью (1), то параметры $a_1, a_2, \ldots, a_m$ могут быть найдены из системы уравнений

$$y_j=\tilde{f}(x_j;\ a_1,\ a_2,\ \ldots,\ a_m) \quad (j=1,\ 2,\ \ldots,\ n). \tag{2}$$

Однако на практике значения (x_i, y_i) содержат неизбежные ошибки и число уравнений системы (2) значительно больше числа неизвестных; поэтому система (2), как правило, является несовместной. Приходится отыскивать наилучшие значения $\tilde{a}_1$, $\tilde{a}_2, \ldots, \tilde{a}_m$, приближенно удовлетворяющие системе (2), т. е. такие, что *невязки (уклонения)*

$$y_j-\tilde{f}(x_j;\ \tilde{a}_1,\ \tilde{a}_2,\ \ldots,\ \tilde{a}_m)=\varepsilon_j \quad (j=1,\ 2,\ \ldots,\ n) \tag{3}$$

являются возможно малыми по абсолютной величине.

Геометрически задача сводится к проведению кривой вида (1), наиболее тесно примыкающей к данной системе точек.

Наиболее распространенными являются эмпирические формулы, линейно зависящие от параметров, т. е. формулы вида

$$y=\varphi_0(x)+a_1\varphi_1(x)+a_2\varphi_2(x)+\ldots+a_m\varphi_m(x). \tag{4}$$

В этом случае система (2) линейная и исследование ее сравнительно просто. При нелинейной зависимости в (1) от параметров $a_1, a_2, \ldots, a_m$ система (2) также нелинейная и нахождение точных или приближенных решений ее представляет трудную задачу; обычно такую систему приближенно заменяют линейной (см. § 12).

В следующих параграфах мы рассмотрим три наиболее употребительных метода определения параметров эмпирической формулы: 1) метод выбранных точек, 2) метод средних и 3) метод наименьших квадратов.

§ 6. Метод выбранных точек

Пусть для системы опытных данных $M_i(x_i, y_i)$ $(i=1, 2, \ldots, n)$ построена эмпирическая формула

$$y=\tilde{f}(x;\ a_1,\ a_2,\ \ldots,\ a_m), \tag{1}$$

содержащая m $(m < n)$ свободных параметров $a_1, a_2, \ldots, a_m$, где $\tilde{f}$ — известная функция.

На координатной плоскости Oxy с возможной аккуратностью проводим плавную кривую Г, наиболее близко примыкающую к точкам M_i. На кривой Г выбираем систему m (по числу параметров) точек $N_j(\tilde{x}_j, \tilde{y}_j)$ $(j=1, 2, \ldots, m)$, не обязательно совпадающих с точками M_i. При этом желательно [3], чтобы выбранные точки N_j были по возможности равномерно распределены по всей рабочей части кривой Г и возможно дальше отстояли друг от друга, а в то же время не лежали бы слишком близко к мало надежным концевым точкам M_1 и M_n. Для удобства обычно берут абсциссы $\tilde{x}_j$ этих точек совпадающими с крупными делениями оси Ox координатной сетки. После этого со всей тщательностью замеряют координаты $\tilde{x}_j$, $\tilde{y}_j$ $(j=1, 2, \ldots, n)$. Тогда параметры $a_1, a_2, \ldots, a_m$ в общем случае могут быть определены из системы m уравнений

$$\tilde{y}_j = \tilde{f}(\tilde{x}_j;\ a_1, a_2, \ldots, a_m) \quad (j=1, 2, \ldots, m). \tag{2}$$

Применение этого метода для линейной зависимости $y = ax + b$ рассмотрено выше (§ 2).

Для случая квадратичной зависимости $y = ax^2 + bx + c$ коэффициенты a, b и c определяются из системы трех уравнений

$$\tilde{y}_1 = a\tilde{x}_1^2 + b\tilde{x}_1 + c, \qquad \tilde{y}_2 = a\tilde{x}_2^2 + b\tilde{x}_2 + c, \qquad \tilde{y}_3 = a\tilde{x}_3^2 + b\tilde{x}_3 + c.$$

Заметим, что метод выбранных точек содержит геометрические построения, допускающие известный произвол, и поэтому является грубым. К нему следует прибегать в тех случаях, когда точность исходных данных относительно невелика. Для увеличения точности метода рекомендуется пользоваться сеткой с мелкими делениями. Достоинство метода — простота применения и наглядность.

В дальнейших параграфах (§ 7 и 8) укажем более точные аналитические методы определения коэффициентов эмпирической формулы (1).

§ 7. Метод средних

Если в эмпирическую формулу

$$y = \tilde{f}(x;\ a_1, a_2, \ldots, a_m) \tag{1}$$

подставить исходные данные $M_i(x_i, y_i)$, то левая часть формулы, вообще говоря, не будет равна правой. Разности (невязки)

$$\tilde{f}(x_i;\ a_1, \ldots, a_m) - y_i = \varepsilon_i \quad (i=1, 2, \ldots, n) \tag{2}$$

называются *уклонениями* и представляют собой расстояния по вертикали точек M_i от графика эмпирической функции (1), взятые со знаком плюс (+) или со знаком минус (—) (рис. 18).

Согласно *методу средних* за наилучшее положение эмпирической кривой K принимается то, для которого равна нулю алгебраическая сумма E всех уклонений ε_i, т. е. должно иметь место равенство

$$E = \sum_{i=1}^{n} \varepsilon_i = 0. \tag{3}$$

Для определения по методу средних постоянных $a_1, a_2, \ldots, a_m$, где $m < n$, все уклонения ε_i разбивают на m групп, содержащих

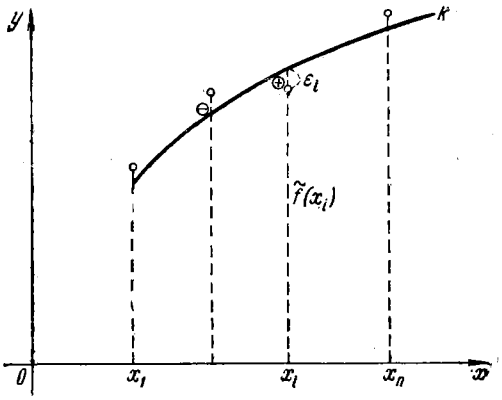

Рис. 18.

примерно одинаковые количества уклонений. Приравнивая нулю алгебраическую сумму E_j $(j = 1, 2, \ldots, m)$ уклонений, входящих в каждую из этих групп, получаем систему, содержащую столько уравнений, сколько имеется неизвестных коэффициентов $a_1, a_2, \ldots, a_m$.

Решив эту систему, мы найдем коэффициенты a_i $(i = 1, 2, \ldots, m)$. Заметим, что так как сумма E_j уклонений для каждой группы равна нулю, то равна нулю также и сумма E всех уклонений, т. е. для нашей системы равенство (3) будет выполнено.

Заметим, что результаты метода средних существенно зависят от способа группировки уклонений. Практика показывает, что наиболее удачные эмпирические формулы получаются, если уклонения группируются в порядке последовательности их номеров (предполагаем, что исходные данные упорядочены по одной из переменных) и каждая группа уклонений содержит по возможности одинаковое число членов.

Пример 1. Количество Q вещества в %, оставшегося в системе через t минут от начала химической реакции, дается таблицей [13]:

t	7	12	17	22	27	32	37
Q	83,7	72,9	63,2	54,7	47,5	41,4	36,3

Составить эмпирическую формулу для зависимости величины Q от времени t.

Решение. На плоскости OtQ данная система точек (t_i, Q_i) примерно располагается на параболе с вертикальной осью. Поэтому будем искать эмпирическую формулу в виде

$$Q = at^2 + bt + c. \tag{4}$$

Для определения коэффициентов a, b и c применим метод средних. Подставляя табличные данные в формулу (4), получаем выражения для соответствующих уклонений:

$$\varepsilon_1 = 49a + 7b + c - 83{,}7; \qquad \varepsilon_5 = 729a + 27b + c - 47{,}5;$$
$$\varepsilon_2 = 144a + 12b + c - 72{,}9; \qquad \varepsilon_6 = 1024a + 32b + c - 41{,}4;$$
$$\varepsilon_3 = 289a + 17b + c - 63{,}2; \qquad \varepsilon_7 = 1369a + 37b + c - 36{,}3.$$
$$\varepsilon_4 = 484a + 22b + c - 54{,}7;$$

Для определения коэффициентов a, b и c по методу средних уклонения ε_i нужно разбить на три группы. Объединим, например, в группу I уклонения ε_1, ε_2, ε_3, в группу II — ε_4, ε_5 и в группу III — ε_6, ε_7. Тогда получим систему

$$\left.\begin{aligned} \varepsilon_1 + \varepsilon_2 + \varepsilon_3 &= 0, \\ \varepsilon_4 + \varepsilon_5 &= 0, \\ \varepsilon_6 + \varepsilon_7 &= 0 \end{aligned}\right\} \tag{5}$$

или

$$\left.\begin{aligned} 482a + 36b + 3c &= 219{,}8, \\ 1213a + 49b + 2c &= 102{,}2, \\ 2393a + 69b + 2c &= 77{,}7. \end{aligned}\right\} \tag{6}$$

Решая систему (6), находим $a = 0{,}0235$; $b = -2{,}6115$; $c = 100{,}8295$. Следовательно, искомая эмпирическая формула имеет вид

$$\tilde{Q} = 0{,}0235t^2 - 2{,}6115t + 100{,}8295. \tag{7}$$

Для сравнения в таблице 15 приведены результаты расчетов по формуле (7).

Таблица 15

Оценка точности эмпирической формулы (7)

t	Q_i	$\tilde{Q}_i$, вычисл. по эмпирической формуле (7)	Уклонения $\varepsilon_i = Q_i - \tilde{Q}_i$
1	83,7	83,70	0,00
2	72,9	72,88	+ 0,02
3	63,2	63,23	— 0,03
4	54,7	54,75	— 0,05
5	47,5	47,45	+ 0,05
6	41,4	41,33	+ 0,07
7	36,3	36,38	+ 0,08

§ 8. Метод наименьших квадратов

Пусть известен вид эмпирической формулы

$$y=\tilde{f}(x;\ a_1,\ a_2,\ \ldots,\ a_m) \tag{1}$$

и

$$\varepsilon_i=\tilde{f}(x_i;\ a_1,\ a_2,\ \ldots,\ a_m)-y_i \qquad (i=1,\ 2,\ \ldots,\ n) \tag{2}$$

— уклонения эмпирической формулы (1) от исходных данных $(x_i,\ y_i)$.

Согласно методу наименьших квадратов (см. гл. I, § 4) наилучшими коэффициентами $a_1,\ a_2,\ \ldots,\ a_m$ считаются те, для которых сумма квадратов уклонений

$$S(a_1,\ a_2,\ \ldots,\ a_m)=\sum_{i=1}^{n}[\tilde{f}(x_i;\ a_1,\ a_2,\ \ldots,\ a_m)-y_i]^2 \tag{3}$$

будет минимальной. Отсюда, используя необходимые условия экстремума функции нескольких переменных, получаем так называемую *нормальную систему* для определения коэффициентов $a_i\ (i=1,\ 2,\ \ldots,\ m)$:

$$\frac{\partial S}{\partial a_1}=0,\quad \frac{\partial S}{\partial a_2}=0,\quad \ldots,\quad \frac{\partial S}{\partial a_m}=0. \tag{4}$$

Если система (4) имеет единственное решение, то оно будет искомым. Система (4) упрощается, если эмпирическая функция $\tilde{f}(x;\ a_1,\ a_2,\ \ldots,\ a_m)$ линейная относительно параметров $a_1,\ a_2,\ \ldots,\ a_m$. Действительно, полагая

$$\tilde{f}(x;\ a_1,\ a_2,\ \ldots,\ a_m)=\varphi_0(x)+a_1\varphi_1(x)+\ldots+a_m\varphi_m(x),$$

будем иметь

$$S(a_1,\ a_2,\ \ldots,\ a_m)=\sum_{i=1}^{n}[a_1\varphi_1(x_i)+\ldots+a_m\varphi_m(x_i)-Y_i]^2,$$

где $Y_i=y_i-\varphi_0(x_i)$. Отсюда

$$\left.\begin{aligned}
&\frac{1}{2}\frac{\partial S}{\partial a_1}=\sum_{i=1}^{n}\varphi_1(x_i)[a_1\varphi_1(x_i)+\ldots+a_m\varphi_m(x_i)-Y_i]=0,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&\frac{1}{2}\frac{\partial S}{\partial a_m}=\sum_{i=1}^{n}\varphi_m(x_i)[a_1\varphi_1(x_i)+\ldots+a_m\varphi_m(x_i)-Y_i]=0.
\end{aligned}\right\} \tag{5}$$

Введя сокращенные обозначения

$$(\varphi_j,\ \varphi_k)=\sum_{i=1}^{n}\varphi_j(x_i)\varphi_k(x_i) \quad \text{и} \quad (\varphi_j,\ Y)=\sum_{i=1}^{n}\varphi_j(x_i)Y_i,$$

систему (5) можно записать в виде *нормальной системы*

$$\left.\begin{array}{l} a_1(\varphi_1, \varphi_1)+a_2(\varphi_1, \varphi_2)+\ldots+a_m(\varphi_1, \varphi_m)=(\varphi_1, Y), \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_1(\varphi_m, \varphi_1)+a_2(\varphi_m, \varphi_2)+\ldots+a_m(\varphi_m, \varphi_m)=(\varphi_m, Y). \end{array}\right\} \quad (5')$$

В частном случае, если эмпирическая функция есть полином*) $y=a_0+a_1x+\ldots+a_mx^m$, то $\varphi_j(x)=x^j$ $(j=0, 1, 2, \ldots, m)$. Отсюда имеем

$$(\varphi_j, \varphi_k)=\sum_{i=1}^{n} x_i^{j+k}=[x^{j+k}], \qquad (\varphi_j, Y)=(\varphi_j, y)=\sum_{i=1}^{n} x_i^j y_i=[x^j y].$$

Следовательно, нормальная система (5′) будет иметь вид

$$\left.\begin{array}{r} a_0n+a_1[x]+a_2[x^2]+\ldots+a_m[x^m]=[y], \\ a_0[x]+a_1[x^2]+a_2[x^3]+\ldots+a_m[x^{m+1}]=[xy], \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_0[x^m]+a_1[x^{m+1}]+a_2[x^{m+2}]+\ldots+a_m[x^{2m}]=[x^m y]. \end{array}\right\} \quad (6)$$

Метод наименьших квадратов обладает тем преимуществом, что если сумма S квадратов уклонений ε_i мала, то сами эти уклонения также малы по абсолютной величине. Для метода средних, где составляется алгебраическая сумма уклонений, такого вывода сделать нельзя.

Недостатком метода наименьших квадратов является громоздкость вычислений. Поэтому к нему прибегают обычно при обработке наблюдений высокой точности, когда нужно получить также весьма точные значения параметров. Заметим, что в этом случае промежуточные вычисления нужно проводить с надлежащим количеством десятичных знаков, так как в противном случае при неблагоприятных условиях искомые коэффициенты будут иметь мало верных знаков. В частности, если происходит потеря цифр при вычитании, то вычисления должны быть проведены с достаточным количеством запасных верных значащих цифр. Здесь следует руководствоваться следующим правилом: если численные значения коэффициентов желательно иметь с m верными значащими цифрами и если предварительные вычисления показывают, что первые p цифр исчезнут при вычитании, то вычисления должны быть произведены с $m+p+1$ верными значащими цифрами на всех стадиях работы [5].

Но грубые значения этих коэффициентов могут быть получены значительно проще, т. е. применение метода не будет оправдано.

---

*) Для удобства обозначений мы изменяем нумерацию коэффициентов a_i.

Пример 1. Используя метод наименьших квадратов, вывести эмпирическую формулу для табличной функции $Q=f(t)$ примера из § 7.

Решение. Полагаем

$$y=ax^2+bx+c. \tag{7}$$

Для вычисления коэффициентов нормальной системы составляем таблицу 16. Тогда имеем следующую систему нормальных уравнений:

$$\left.\begin{aligned} 4088a+\ \ 154b+\ \ 7c&=\ \ 399{,}7,\\ 120736a+\ \ 4088b+\ 154c&=\ \ 7688{,}9,\\ 3795092a+120736b+4088c&=186054{,}3. \end{aligned}\right\}$$

Решив эту систему, получим $a=0{,}023381$; $b=-2{,}6066$; $c=100{,}791$.

Таблица 16

Определение параметров эмпирической формулы (7) методом наименьших квадратов

t^0	t	t^2	t^3	t^4	Q	tQ	t^2Q
1	7	49	343	2401	83,7	585,9	4101,3
1	12	144	1728	20736	72,9	874,8	10497,6
1	17	289	4913	83521	63,2	1074,4	18264,8
1	22	484	10648	234256	54,7	1203,4	26474,8
1	27	729	19683	531441	47,5	1282,5	34627,5
1	32	1024	32768	1048576	41,4	1324,8	42393,6
1	37	1369	50653	1874161	36,3	1343,1	49694,7
7	154	4088	120736	3795092	399,7	7688,9	186054,3

Следовательно, искомая эмпирическая формула запишется так:

$$\tilde{Q}=0{,}02338t^2-2{,}6066t+100{,}79. \tag{8}$$

Таблица 17 показывает согласованность полученной эмпирической формулы с опытными данными.

Таблица 17

Оценка точности эмпирической формулы (8)

i	Q	$\tilde{Q}$, вычисл. по формуле (8)	Уклонение $\varepsilon=Q-\tilde{Q}$
1	83,7	83,69	+0,01
2	72,9	72,88	+0,02
3	63,2	63,24	—0,04
4	54,7	54,76	—0,06
5	47,5	47,46	+0,04
6	41,4	41,32	+0,08
7	36,3	36,36	—0,06

Имеем $\sum_{i=1}^{7} \varepsilon_i^2 = 0{,}0173$.

Мы видим, что формула (8) согласуется с экспериментальными данными несколько лучше, чем формула (7) из § 7, вычисленная методом средних.

Пример 2. Следующая таблица [1]

t	14,5	30,0	64,5	74,5	86,7	94,5	98,9
k	0	0,004	0,018	0,029	0,051	0,073	0,090

дает значения удельной электропроводимости k стекла в зависимости от температуры t в градусах С. Подобрать эмпирическую формулу для функции $k = f(t)$.

Решение. Все точки $M_i(t_i, \lg k_i)$ $(i = 1, 2, \ldots, 7)$, за исключением первой, примерно расположены на прямой линии (рис. 19).

Поэтому выбираем эмпирическую формулу в виде показательной функции

$$k = be^{at}.$$

Для удобства положим

$$1000k = ce^{at}, \qquad (9)$$

где $c = 1000b$. Логарифмируя (9), будем иметь

$$y = \lg 1000k = \lg c + aMt, \qquad (10)$$

где $M = \lg e = 0{,}43429$.

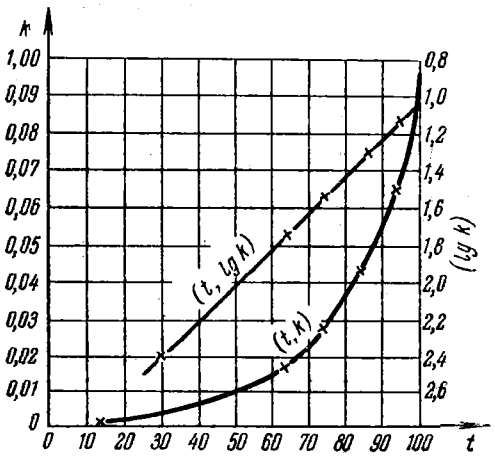

Рис. 19.

Преобразованные данные помещены в таблице 18.

Таблица 18

Значения переменных t и y

t	$1000k$	$y = \lg 1000k$	k_{I}	$\Delta_{\text{I}} = k - k_{\text{I}}$	k_{II}	$\Delta_{\text{II}} = k - k_{\text{II}}$
14,5	0	—	0,0019	—0,0019	0,0025	—0,0025
30,0	4	0,6021	0,0039	—0,0001	0,0048	—0,0008
64,5	18	1,2553	0,0185	—0,0005	0,0200	—0,0020
74,5	29	1,4624	0,0293	—0,0003	0,0302	—0,0012
86,7	51	1,7076	0,0512	—0,0002	0,0499	+0,0011
94,5	73	1,8633	0,0731	—0,0001	0,0688	+0,0042
98,9	90	1,9542	0,0893	—0,0007	0,0825	+0,0075

оффициенты $\lg c$ и aM в формуле (10) определим двумя способами, ориируя первую точку.

I. Метод средних. Определяем уклонения:

$$\varepsilon_1 = 0{,}6021 - (\lg c + 30aM); \qquad \varepsilon_4 = 1{,}7076 - (\lg c + 86{,}7aM);$$
$$\varepsilon_2 = 1{,}2553 - (\lg c + 64{,}5aM); \qquad \varepsilon_5 = 1{,}8633 - (\lg c + 94{,}5aM);$$
$$\varepsilon_3 = 1{,}4624 - (\lg c + 74{,}5aM); \qquad \varepsilon_6 = 1{,}9542 - (\lg c + 98{,}9aM).$$

Отсюда, полагая $\varepsilon_1 + \varepsilon_2 + \varepsilon_3 = 0$, $\varepsilon_4 + \varepsilon_5 + \varepsilon_6 = 0$, получим систему

$$3 \lg c + 169aM = 3{,}3198, \qquad 3 \lg c + 280{,}1aM = 5{,}5251. \tag{11}$$

Решая систему (11), находим

$$aM = \frac{2{,}2053}{111{,}1} = 0{,}01985,$$

$$a = \frac{1}{M} \cdot 0{,}01985 = 2{,}30258 \cdot 0{,}01985 = 0{,}0457$$

и

$$\lg c = \frac{1}{3}(3{,}3198 - 169 \cdot 0{,}01985) = -0{,}0116 = \bar{1}{,}9884,$$

$$c = 0{,}9737.$$

Следовательно,

$$k = 0{,}9737 \cdot 10^{-3} \cdot e^{0{,}0457t}. \tag{12}$$

II. Метод наименьших квадратов. Промежуточные вычисления приведены в таблице 19.

Таблица 19

Определение параметров формулы (10) методом наименьших квадратов

t^0	t	t^2	y	ty
1	30,0	900	0,6021	18,063
1	64,5	4160,25	1,2553	80,967
1	74,5	5550,25	1,4624	108,949
1	86,7	7516,89	1,7076	148,049
1	94,5	8930,25	1,8633	176,082
1	98,9	9781,21	1,9542	193,70
6	449,1	36838,85	8,8449	725,380

Отсюда получаем нормальную систему

$$6 \lg c + 449{,}1aM = 8{,}8449, \qquad 449{,}1 \lg c + 36\,839aM = 725{,}380$$

или

$$\lg c + 74{,}183aM = 1{,}4742, \qquad \lg c + 82{,}026aM = 1{,}6145.$$

Решая последнюю систему, находим $aM=0{,}01789$; $a=0{,}0412$ и $\lg c=0{,}1471$; $c=1{,}403$. Таким образом,

$$k=1{,}403\cdot 10^{-3}\cdot e^{0{,}0412t}. \tag{13}$$

Результаты вычислений по формулам (12) и (13) приведены в таблице 18.

§ 9. Некоторые соображения о выборе вида эмпирической формулы с двумя параметрами

Пусть для данной системы значений x_i, y_i $(i=1, 2, \ldots, n)$, где $n\geqslant 3$ и $x_1<x_2<\ldots<x_n$, требуется найти эмпирическую формулу вида

$$y=\tilde{f}(x;\ a,\ b), \tag{1}$$

содержащую лишь два параметра a и b.

Если окажется, что

$$\frac{\Delta y_i}{\Delta x_i}=\text{const}, \tag{2}$$

то искомая зависимость линейная:

$$y=ax+b, \tag{3}$$

и задача, таким образом, легко решается.

Другим простым случаем является наличие квадратичной зависимости

$$y-y_1=a(x-x_1)^2+b(x-x_1), \tag{4}$$

которая обнаруживается известными приемами (см. § 3).

Рассмотрим общий случай, когда соотношение (1), вообще говоря, не сводится к формулам (3) и (4).

Достаточным условием существования эмпирической формулы вида (1), где $\tilde{f}$—известная функция, является совместность (с заданной точностью) системы уравнений $y_i=\tilde{f}(x_i;\ a,\ b)$ $(i=1, 2, \ldots, n)$. Исключая отсюда неизвестные a и b, получаем систему условий для точек (x_i, y_i), обеспечивающую существование зависимости (1). Но такой подход является весьма сложным.

Выведем необходимое условие существования эмпирической зависимости вида (1) для заданной системы точек (x_i, y_i). Пусть $M_i(x_i, y_i)$, $M_j(x_j, y_j)$, $M_k(x_k, y_k)$—три системы значений из нашей совокупности. Предполагая, что кривая (1) проходит через точки M_i, M_j, M_k, будем иметь:

$$y_i=\tilde{f}(x_i;\ a,\ b),\quad y_j=\tilde{f}(x_j;\ a,\ b),\quad y_k=\tilde{f}(x_k;\ a,\ b). \tag{5}$$

Исключая из системы (5) параметры a и b, получим соотношение вида

$$\Phi(x_i,\ x_j,\ x_k,\ y_i,\ y_j,\ y_k)=0. \tag{6}$$

Выполнение равенства (6) для любых значений i, j, k ($1 \leqslant i < j < k \leqslant n$) необходимо для существования зависимости (1).

Так как проверка соотношения (6) связана с трудоемкими вычислениями, то на практике обычно ограничиваются одной тройкой точек: начальной (x_1, y_1), промежуточной (x_s, y_s) и конечной (x_n, y_n) (с целью достижения наибольшего диапазона), т. е. полагают $i=1$, $j=s$ $(1 < s < n)$, $k=n$.

Точку M_s выбирают так, чтобы соотношение (6) было бы по возможности простым. Заметим, что иногда вместо промежуточной точки M_s выгодно брать точку $\tilde{M}_s(\tilde{x}_s, \tilde{y}_s)$, не принадлежащую нашему ряду точек M_i; тогда координаты $\tilde{x}_s$ и $\tilde{y}_s$ определяются интерполированием.

Пример 1. Получить необходимое условие для существования степенной зависимости

$$y = ax^b, \tag{7}$$

предполагая, что $x_i > 0$, $y_i > 0$ $(i=1, 2, \ldots, n)$.

Решение. Выберем

$$x_s = \sqrt{x_1 x_n}.$$

Из формулы (7) имеем

$$y_1 = ax_1^b, \quad y_s = ax_s^b = a x_1^{\frac{b}{2}} x_n^{\frac{b}{2}}, \quad y_n = ax_n^b. \tag{8}$$

Исключая из соотношений (8) параметры a и b, получим

$$y_1 y_n = y_s^2, \quad \text{т. е.} \quad y_s = \sqrt{y_1 y_n}. \tag{9}$$

Таким образом, для существования степенной зависимости (7) необходимо, чтобы среднему геометрическому x_s значений x_1 и x_n соответствовало среднее геометрическое y_s значений y_1 и y_n. Вообще, если имеет место степенная зависимость (7) и значения x_i образуют геометрическую прогрессию, то значения y_i также образуют геометрическую прогрессию.

Если значение $x_s = \sqrt{x_1 x_n}$ не является табличным, то соответствующее значение $\tilde{y}_s$ определяется с помощью интерполирования.

В дальнейшем в этом параграфе мы будем рассматривать наиболее часто встречающиеся зависимости:

I. $y = ax + b$. IV. $y = a + \frac{b}{x}$. VI. $y = \frac{x}{ax+b}$.

II. $y = ax^b$. V. $y = \frac{1}{ax+b}$. VII. $y = a \ln x + b$.

III. $y = ab^x$.

Аналогично тому, как это сделано в примере 1, для существования зависимостей I—VII легко вывести простые необходимые условия вида

$$x_s = \bar{x}_s, \quad y_s = \bar{y}_s,$$

где $\bar{x}_s = \varphi(x_1, x_n)$ и $\bar{y}_s = \psi(y_1, y_n)$, причем предполагается, что $x_i > 0$ и $y_i > 0$ $(i = 1, 2, \ldots, n)$. Выражения для $\bar{x}_s$ и $\bar{y}_s$ приведены в таблице 20.

Таблица 20

Простейшие необходимые условия для наличия эмпирических зависимостей I — VII

№	$\bar{x}_s$	$\bar{y}_s$	Вид эмпирической формулы	Способ выравнивания
I	$\frac{x_1+x_n}{2}$ (среднее арифметическое)	$\frac{y_1+y_n}{2}$ (среднее арифметическое)	$y = ax + b$	
II	$\sqrt{x_1 x_n}$ (среднее геометрическое)	$\sqrt{y_1 y_n}$ (среднее геометрическое)	$y = ax^b$	$Y = \alpha + bX$, где $X = \lg x$, $Y = \lg y, \alpha = \lg a$
III	$\frac{x_1+x_n}{2}$ (среднее арифметическое)	$\sqrt{y_1 y_n}$ (среднее геометрическое)	$y = ab^x$ или $y = ae^{\beta x}$, где $\beta = \ln b$	$Y = \alpha + \beta x$, где $Y = \lg y$, $\alpha = \lg a$, $\beta = \lg b$
IV	$\frac{2x_1x_n}{x_1+x_n}$ (среднее гармоническое)	$\frac{y_1+y_n}{2}$ (среднее арифметическое)	$y = a + \frac{b}{x}$	$Y = ax + b$, где $Y = xy$
V	$\frac{x_1+x_n}{2}$ (среднее арифметическое)	$\frac{2y_1y_n}{y_1+y_n}$ (среднее гармоническое)	$y = \frac{1}{ax+b}$	$Y = ax + b$, где $Y = \frac{1}{y}$
VI	$\frac{2x_1x_n}{x_1+x_n}$ (среднее гармоническое)	$\frac{2y_1y_n}{y_1+y_n}$ (среднее гармоническое)	$y = \frac{x}{ax+b}$	$Y = ax + b$, где $Y = \frac{x}{y}$
VII	$\sqrt{x_1 x_n}$ (среднее геометрическое)	$\frac{y_1+y_n}{2}$ (среднее арифметическое)	$y = a \lg x + b$	$y = aX + b$, где $X = \lg x$

Таблица 20 облегчает выбор вида эмпирической формулы среди указанных. Рекомендуется поступать следующим образом: для проверки пригодности определенной эмпирической формулы, пользуясь исходными данными, находим значения $x_s = \bar{x}_s$ и y_s и сравниваем последнее со значением $\psi(y_1, y_n) = \bar{y}_s$, помещенным в таблице. Предпочтительнее та эмпирическая формула, для которой расхождение $|y_s - \bar{y}_s|$ возможно мало. Для окончательного выбора следует учесть также промежуточные данные. Если величина $|y_s - \bar{y}_s|$ большая, то соответствующая эмпирическая формула непригодна.

Если значение $\varphi(x_1, \bar{x}_n) = x_s$ не находится среди исходных данных x_i, то отвечающее ей значение можно определить посредством линейной интерполяции

$$\hat{y}_s = y_i + \frac{y_{i+1} - y_i}{x_{i+1} - x_i}(\bar{x}_s - x_i),$$

где x_i и x_{i+1} — промежуточные значения, между которыми содержится $\bar{x}_s$ $(x_i < \bar{x}_s < x_{i+1})$.

Следует иметь в виду, что такой подход в целом является грубо ориентировочным, так как мы не учитываем поведение всех промежуточных данных (x_i, y_i). Кроме того, таблица 20 охватывает небольшое количество зависимостей и может случиться, что переменные x и y подчиняются некоторой закономерности, не вошедшей в наш список.

Заметим, что все зависимости, приведенные в таблице 20, методом выравнивания могут быть просто преобразованы в линейные (ср. § 2). Поэтому здесь можно использовать также критерий прямолинейности (§ 2, формула (6)) для преобразованных исходных данных (x_i, y_i).

Функции I—VII монотонные, и, следовательно, отвечающие им упорядоченные данные (x_i, y_i) при $\Delta x_i = x_{i+1} - x_i > 0$ $(i = 1, 2, \ldots, n-1)$ должны обладать постоянным знаком приращения $\Delta y_i = y_{i+1} - y_i$ $(i = 1, 2, \ldots, n-1)$.

Если это обстоятельство не имеет места, то зависимости I—VII противопоказаны.

Пример 2. Определить вид эмпирической формулы, отвечающей следующей таблице:

x	273	283	288	293	313	333	353	373
y	29,4	33,3	35,2	37,2	45,8	55,2	65,6	77,3

Таблица 21

Подбор эмпирической формулы на основании критериев I — VII (см. табл. 20)

	$\bar{x}_s$	$\bar{y}_s$	$\hat{y}_s$	$\lvert\hat{y}_s-\bar{y}_s\rvert$	Вид формулы
I	$\frac{x_1+x_n}{2}=\frac{273+373}{2}=323$	$\frac{y_1+y_n}{2}=\frac{29,4+77,3}{2}=53,35$	50,5	2,85	$y=ax+b$ — мало подходит
II	$\sqrt{x_1x_n}=\sqrt{273\cdot 373}=319,1$	$\sqrt{y_1y_n}=\sqrt{29,4\cdot 77,3}=47,7$	48,7	1,0	$y=ax^b$ — подходит лучше других формул
III	$\frac{x_1+x_n}{2}=323$	$\sqrt{y_1y_n}=47,7$	50,5	2,8	$y=ab^x$ — мало подходит
IV	$\frac{2x_1x_n}{x_1+x_n}=2\cdot\frac{273\cdot 373}{273+373}=315,3$	$\frac{y_1+y_n}{2}=53,35$	46,9	6,45	$y=a+\frac{b}{x}$ — не подходит
V	$\frac{x_1+x_n}{2}=323$	$\frac{2y_1y_n}{y_1+y_n}=2\cdot\frac{29,4\cdot 77,3}{29,4+77,3}=42,6$	50,5	7,9	$y=\frac{1}{ax+b}$ — не подходит
VI	$\frac{2x_1x_n}{x_1+x_n}=315,3$	$\frac{2y_1y_n}{y_1+y_n}=42,6$	46,9	4,3	$y=\frac{x}{ax+b}$ — не подходит
VII	$\sqrt{x_1x_n}=319,1$	$\frac{y_1+y_n}{2}=53,35$	48,7	4,65	$y=a\lg x+b$ — не подходит

Решение. Будем искать эмпирическую формулу среди зависимостей I—VII, имеющихся в таблице 20, согласно указанному рецепту. Результаты вычислений приведены в таблице 21, причем в необходимых случаях применена линейная интерполяция.

Таблица 21 показывает, что согласно необходимому критерию следует остановиться на степенной зависимости $y=ax^b$.

Пример 3. Следующая таблица дает давление p насыщенного пара в $кГ/см^2$, соответствующее удельному объему v в $м^3/кГ$.

v	3,334	1,630	0,8657	0,4323	0,2646	0,1699	0,1146
p	0,482	1,034	2,027	4,247	7,164	11,48	17,60

Найти эмпирическую формулу для зависимости $F(v, p)=0$ [1].

Решение. Подберем вид формулы, пользуясь таблицей 20. Остановившись на формуле II, имеем

$$\bar{v}_s=\sqrt{v_1 v_n}=\sqrt{3,334\cdot 0,1146}=0,618,$$
$$\bar{p}_s=\sqrt{p_1 p_n}=\sqrt{0,482\cdot 17,60}=2,92.$$

Значение $\hat{p}_s$, соответствующее $\hat{v}_s=0,618$, найдем линейной интерполяцией:

$$\hat{p}_s=2,027+\frac{4,247-2,027}{0,4323-0,8657}\cdot(0,618-0,8657)=3,295.$$

Так как отклонение $|\bar{p}_s-\hat{p}_s|=0,375$ сравнительно незначительное, то можно выбрать эмпирическую формулу вида

$$p=av^b. \tag{10}$$

Это согласуется также с природой газовых законов

Логарифмируя формулу (10), будем иметь

$$\lg p=\lg a+b\lg v.$$

Отсюда, полагая

$$X=\lg v,\quad Y=\lg p,$$

получим линейную зависимость

$$Y=bX+\lg a. \tag{11}$$

Значения

$$X_i=\lg v_i \text{ и } Y_i=\lg p_i \quad (i=1, 2, \ldots, 7)$$

приведены в таблице 22.

Построив точки $M_i(\lg v_i, \lg p_i)$*), убеждаемся, что они расположены на прямой линии (рис. 20); следовательно, выбор эмпирической формулы произведен правильно.

*) При построении рекомендуется пользоваться логарифмической бумагой.

Коэффициенты формулы (11) вычисляем методом средних. Составляя уклонения $\varepsilon_i = \lg a + b \lg v_i - \lg p_i$ $(i = 1, 2, \ldots, 7)$ и разбивая их на две группы:

$$\left.\begin{aligned} \varepsilon_1 + \varepsilon_2 + \varepsilon_3 = 0, \\ \varepsilon_4 + \varepsilon_5 + \varepsilon_6 + \varepsilon_7 = 0, \end{aligned}\right\}$$

получим систему

$$\left.\begin{aligned} 3 \lg a + 0{,}6725 b = 0{,}0043, \\ 4 \lg a - 2{,}6522 b = 3{,}7886, \end{aligned}\right\}$$

откуда

$$b = -1{,}066,$$
$$\lg a = 0{,}2402$$

и

$$a = 1{,}740.$$

Следовательно, окончательно имеем

$$p v^{1{,}066} = 1{,}740. \qquad (12)$$

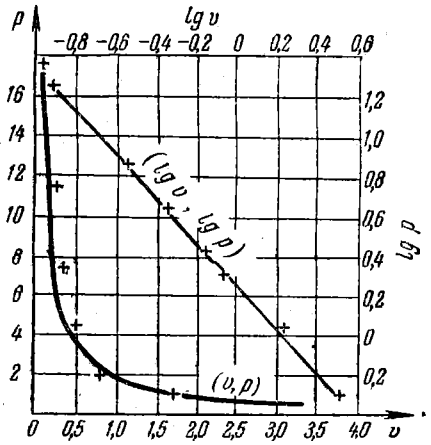

Рис. 20.

В таблице 22 даны расхождения значений $p_э$, полученных по формуле (12), с табличными данными p.

Таблица 22

Оценка точности эмпирической формулы (12)

v	p	$\lg v$	$\lg p$	$p_э$, вычислен. по ф-ле (12)	$\Delta = p - p_э$	Δ^2
3,334	0,482	0,5229	—0,3170	0,482	0,000	0
1,630	1,034	0,2122	0,0145	1,033	0,001	$1 \cdot 10^{-6}$
0,8657	2,027	—0,0626	0,3068	2,028	—0,001	$1 \cdot 10^{-6}$
0,4323	4,247	—0,3642	0,6281	4,241	0,006	$36 \cdot 10^{-6}$
0,2646	7,164	—0,5774	0,8551	7,176	—0,012	$144 \cdot 10^{-6}$
0,1699	11,48	—0,7698	1,0599	11,48	0,00	0
0,1146	17,60	—0,9408	1,2455	17,51	0,090	$8100 \cdot 10^{-6}$
					$\sum = 0{,}084$	$\sum = 0{,}0083$

§ 10. Эмпирические формулы, содержащие три параметра

В этом параграфе мы рассмотрим, для заданной системы значений (x_i, y_i) $(i = 1, 2, \ldots, n)$ табличных данных, важнейшие представители эмпирических формул вида

$$y = \tilde{f}(x;\ a,\ b,\ c), \qquad (1)$$

где a, b и c — некоторые постоянные.

1° Квадратичная зависимость. Пусть

$$y = ax^2 + bx + c. \tag{2}$$

Критерии для квадратичной зависимости были указаны в § 4.

2° Степенная зависимость. Положим

$$y = ax^b + c. \tag{3}$$

Отсюда $y - c = ax^b$. Логарифмируя это выражение, будем иметь

$$\lg(y - c) = \lg a + b \lg x.$$

Отсюда, полагая $\lg(y - c) = Y$ и $\lg x = X$, получим линейную зависимость

$$Y = bX + \lg a. \tag{4}$$

Определение параметров формулы (3) следует начать с нахождения c. Для этого составим среднее геометрическое $x_s = \sqrt{x_1 x_n}$, где x_1 и x_n — крайние значения переменной x, и, пользуясь чертежом или методом линейной интерполяции для x_s, найдем соответствующее значение y_s. Предполагая, что точки $M_1(x_1, y_1)$, $M_s(x_s, y_s)$, $M_n(x_n, y_n)$ расположены на кривой (3), будем иметь три равенства:

$$y_1 = c + ax_1^b, \quad y_s = c + ax_s^b, \quad y_n = c + ax_n^b.$$

Возводя $x_s = \sqrt{x_1 x_n}$ в степень b и умножая на a, получим

$$ax_s^b = \sqrt{ax_1^b ax_n^b} \quad \text{или} \quad y_s - c = \sqrt{(y_1 - c)(y_n - c)}.$$

Решая последнее равенство относительно c, находим

$$c = \frac{y_1 y_n - y_s^2}{y_1 + y_n - 2y_s}.$$

Когда c определено, строим точки $N_i(X_i, Y_i)$, где $X_i = \lg x_i$, $Y_i = \lg(y_i - c)$ $(i = 1, 2, \ldots, n)$. Если эти точки располагаются прямолинейно (или почти прямолинейно), то оправдана зависимость (3), причем постоянные a и b находятся обычным способом.

Пример 1. Для переменных x и y дана таблица значений:

x	250	500	900	1200	1600	2000
y	0,10	0,28	0,80	1,38	2,56	4,10

Найти эмпирическую формулу, связывающую эти переменные [12].

Решение. Построим эмпирическую формулу вида

$$y = ax^b + c.$$

Находим $x_s = \sqrt{x_1 x_n} = \sqrt{250 \cdot 2000} = 707$. На графике этому значению x_s

соответствует $y_s = 0{,}507$; отсюда

$$c = \frac{0{,}10 \cdot 4{,}10 - (0{,}507)^2}{0{,}10 + 4{,}10 - 2 \cdot 0{,}507} = 0{,}048.$$

Остальные параметры a и b найдем методом средних.

Составляем начальные уравнения

$$\lg(y_i - 0{,}048) = \lg a + b \lg x_i \quad (i = 1, 2, \ldots, 6).$$

Будем иметь

$$-1{,}2840 = \lg a + 2{,}3979\, b, \quad 0{,}1245 = \lg a + 3{,}0792\, b,$$
$$-0{,}6345 = \lg a + 2{,}6990\, b, \quad 0{,}4000 = \lg a + 3{,}2041\, b,$$
$$-0{,}1238 = \lg a + 2{,}9542\, b, \quad 0{,}6077 = \lg a + 3{,}3010\, b.$$

Группируя эти уравнения по три, получим

$$-2{,}0423 = 3 \lg a + 8{,}0511\, b, \quad 1{,}1322 = 3 \lg a + 9{,}5843\, b. \tag{5}$$

Решая систему (5), находим $a = 5{,}789 \cdot 10^{-7}$ и $b = 2{,}071$. Следовательно, искомая эмпирическая формула будет иметь вид

$$\tilde{y} = 5{,}789 \cdot 10^{-7} x^{2{,}071} + 0{,}048. \tag{6}$$

Сравнение значений $\tilde{y}$, полученных по формуле (6), с табличными данными y показано в таблице 23.

Таблица 23

Оценка точности эмпирической формулы (6)

x	y	$\tilde{y}$	$\varepsilon = y - \tilde{y}$	x	y	$\tilde{y}$	$\varepsilon = y - \tilde{y}$
250	0,10	0,102	−0,002	1200	1,38	1,426	−0,046
500	0,28	0,273	+0,007	1600	2,56	2,552	+0,008
900	0,80	0,808	−0,008	2000	4,10	4,020	+0,080

3° Показательная зависимость. Пусть

$$y = ae^{bx} + c. \tag{7}$$

Перенося слагаемое c влево и логарифмируя, получим

$$\lg(y - c) = \lg a + (bM)\, x,$$

где $M = 0{,}43429$.

Таким образом,

$$Y = \lg a + bMx, \tag{8}$$

где $Y = \lg(y - c)$.

Сначала определим параметр c. Для этого, как и в предыдущем случае, выберем крайние точки $M_1(x_1, y_1)$ и $M_n(x_n, y_n)$ и составим среднее арифметическое

$$x_s = \frac{x_1 + x_n}{2}.$$

Для значения x_s найдем соответствующее значение y_s (или из чертежа,

или линейной интерполяцией). Подставляя эти значения в эмпирическую формулу (7), будем иметь

$$y_1 = ae^{bx_1} + c; \quad y_n = ae^{bx_n} + c; \quad y_s = ae^{\frac{b}{2}(x_1 + x_n)} + c.$$

Отсюда $y_1 - c = ae^{bx_1}$, $y_n - c = ae^{bx_n}$, и, следовательно,

$$(y_1 - c)(y_n - c) = (y_s - c)^2.$$

Решив последнее уравнение относительно c, получим

$$c = \frac{y_1 y_n - y_s^2}{y_1 + y_n - 2y_s}.$$

Если обнаружена линейная зависимость (8), то остальные параметры a и b находятся обычными приемами.

Выведем аналитический критерий для показательной зависимости (7), предполагая, что значения x_i равноотстоящие, т. е.

$$\Delta x_i = h = \text{const} \quad (i = 1, 2, \ldots, n-1).$$

Из формулы (7) имеем $y_i = ae^{bx_i} + c$ и $y_{i+1} = ae^{b(x_i + h)} + c$. Отсюда

$$y_{i+1} - y_i \equiv \Delta y_i = a_1 e^{bx_i}, \tag{9}$$

где $a_1 = a(e^{bh} - 1)$. Логарифмируя равенство (9), получим

$$\lg \Delta y_i = \lg a_1 + bMx_i.$$

Таким образом, при наличии зависимости (7) точки $N_i(x_i, \lg\Delta y_i)$ $(i = 1, 2, \ldots, n-1)$ расположены на прямой линии. Следовательно, учитывая, что Δx_i постоянны, получаем искомый критерий

$$\Delta(\lg \Delta y_i) = \text{const} \quad (i = 1, 2, \ldots, n-2). \tag{10}$$

Замечание. Что касается эмпирических формул, содержащих свыше трех параметров, то они редко встречаются на практике, и мы детально рассматривать их не будем.

Пример 2. Для переменных x и y дана таблица значений:

x	0	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1,0
y	1,30	1,44	1,59	1,78	1,97	2,19	2,46	2,74	3,06	3,42	3,84

Найти эмпирическую формулу для зависимости между x и y.

Решение. Составим таблицу разностей Δy_i (таблица 24).

Таблица 24

Проверка на показательную зависимость

x	y	Δy	$\lg(\Delta y)$	$\Delta(\lg \Delta y)$	x	y	Δy	$\lg(\Delta y)$	$\Delta(\lg \Delta y)$
0	1,30	0,14	—0,854	0,030	0,6	2,46	0,28	—0,553	0,058
0,1	1,44	0,15	—0,824	0,103	0,7	2,74	0,32	—0,495	0,051
0,2	1,59	0,19	—0,721	0,000	0,8	3,06	0,36	—0,444	0,067
0,3	1,78	0,19	—0,721	0,137	0,9	3,42	0,42	—0,377	
0,4	1,97	0,22	—0,658	0,089	1,0	3,84			
0,5	2,19	0,27	—0,569	0,016					

Так как разность Δy монотонно растет, то зависимость между x и y не является линейной.

Проверяем критерий (10) для показательной зависимости (7). Из таблицы 24 видно, что имеется значительный разброс для $\Delta(\lg \Delta y)$. Поэтому формулу $y = ae^{bx} + c$ можно принять лишь в качестве грубого приближения.

Найдем

$$x_s = \frac{1}{2}(x_1 + x_n) = \frac{1}{2}(0 + 1{,}0) = 0{,}5, \qquad y_s = 2{,}19;$$

отсюда

$$c = \frac{3{,}84 \cdot 1{,}30 - (2{,}19)^2}{1{,}30 + 3{,}84 - 2 \cdot 2{,}19} = 0{,}258.$$

Параметры a и b определим по методу средних. Имеем

$$\lg(y_i - 0{,}258) = \lg a + (bM)\, x_i \qquad (i = 0, 1, \ldots, 10).$$

Составим два уравнения:

$$\sum_{i=0}^{5} \lg(y_i - 0{,}258) = 6 \lg a + bM \sum_{i=0}^{5} x_i,$$

$$\sum_{i=6}^{10} \lg(y_i - 0{,}258) = 5 \lg a + bM \sum_{i=6}^{10} x_i.$$

Подставив значения x_i и y_i, получим систему

$$0{,}9169 = 6 \lg a + 1{,}5bM, \qquad 2{,}2392 = 5 \lg a + 4{,}0bM,$$

откуда $a = 1{,}044$; $b = 1{,}234$. Следовательно, искомая эмпирическая формула имеет вид

$$\tilde{y} = 1{,}044e^{1{,}234x} + 0{,}258. \tag{11}$$

Согласованность полученной формулы с исходными данными показана в таблице 25.

Таблица 25

Уклонения эмпирической формулы (11) от табличных данных

x	y	$\tilde{y}$, вычисл. по формуле (11)	$\Delta = y - \tilde{y}$
0	1,30	1,302	−0,002
0,1	1,44	1,439	+0,001
0,2	1,59	1,594	−0,004
0,3	1,78	1,770	+0,010
0,4	1,97	1,968	+0,002
0,5	2,19	2,193	−0,003
0,6	2,46	2,447	+0,013
0,7	2,74	2,735	+0,005
0,8	3,06	3,048	+0,012
0,9	3,42	3,428	−0,008
1,0	3,84	3,844	−0,004

§ 11. Уточнение полученной эмпирической формулы

Укажем простой прием уточнения полученной эмпирической формулы в случае, если она дает сумму квадратов уклонений более значительную, чем это желательно.

Пусть для заданной системы значений (x_i, y_i) $(i=1, 2, \ldots, n)$ найдена эмпирическая формула

$$y=f(x). \tag{1}$$

Требуется уменьшить сумму квадратов уклонений

$$s=\sum_{i=1}^{n}\varepsilon_i^2=\sum_{i=1}^{n}[y_i-f(x_i)]^2, \tag{2}$$

где $\varepsilon_i=y_i-f(x_i)$ $(i=1, 2, \ldots, n)$ *).

Рассмотрим функцию

$$\tilde{y}=f(x)+c,$$

где c — некоторая постоянная величина.

Подберем число c так, чтобы сумма квадратов новых уклонений

$$\tilde{s}=\sum_{i=1}^{n}\tilde{\varepsilon}_i^2=\sum_{i=1}^{n}[y_i-f(x_i)-c]^2 \tag{3}$$

была минимальной. Очевидно, имеем

$$\tilde{s}=\sum_{i=1}^{n}(c-\varepsilon_i)^2.$$

Для минимума функции $\tilde{s}=\tilde{s}(c)$ необходимо, чтобы

$$\frac{\partial\tilde{s}}{\partial c}=2\sum_{i=1}^{n}(c-\varepsilon_i)=0.$$

Отсюда $nc=\sum_{i=1}^{n}\varepsilon_i$, и, следовательно,

$$c=\frac{1}{n}\sum_{i=1}^{n}\varepsilon_i. \tag{4}$$

Так как $\frac{\partial^2\tilde{s}}{\partial c^2}=2n>0$, то значение (4) дает наименьшее значение для суммы квадратов $\tilde{s}$.

Итак, наилучшей постоянной c является среднее арифметическое уклонений ε_i.

---

*) Для упрощения дальнейших формул мы изменяем знак уклонения.

Таким образом, если $c \neq 0$, то прибавление к правой части эмпирической формулы (1) постоянного числа c, определяемого формулой (4), приводит к уточнению этой формулы в смысле суммы квадратов уклонений, так как $\tilde{s} < s$. Если $c = 0$ или близко к нулю, то указанный прием не дает нужного эффекта. В этом случае можно положить

$$\tilde{y} = f(x) + c\varphi(x),$$

где $\varphi(x)$ — известная функция, не обращающаяся в нуль во всех точках x_i, т. е. такая, что $\sum_{i=1}^{n} \varphi(x_i)^2 > 0$. Отсюда

$$\tilde{s} = \sum_{i=1}^{n} [y_i - f(x_i) - c\varphi(x_i)]^2 = \sum_{i=1}^{n} [c\varphi(x_i) - \varepsilon_i]^2.$$

Необходимое условие для минимума функции $\tilde{s}$ дает

$$\frac{1}{2}\frac{\partial \tilde{s}}{\partial c} = \sum_{i=1}^{n} \varphi(x_i)[c\varphi(x_i) - \varepsilon_i] = 0;$$

следовательно,

$$c = \frac{\sum_{i=1}^{n} \varepsilon_i \varphi(x_i)}{\sum_{i=1}^{n} [\varphi(x_i)]^2}. \tag{5}$$

Так как

$$\frac{1}{2}\frac{\partial^2 \tilde{s}}{\partial c^2} = \sum_{i=1}^{n} \varphi^2(x_i) > 0,$$

то наилучшей постоянной для нашей цели является «среднее взвешенное» уклонений ε_i, определяемое формулой (5). В частности, если $\varphi(x) = x$, то имеем

$$c = \frac{\sum_{i=1}^{n} \varepsilon_i x_i}{\sum_{i=1}^{n} x_i^2}.$$

Пример 1. Для табличных данных

x	8	10	15	20	30	40	60	80
y	1,30	14,0	15,4	16,3	17,2	17,8	18,5	18,8

получена эмпирическая формула

$$\bar{y}=\frac{x}{0{,}051x+0{,}209}. \tag{6}$$

Путем прибавления постоянной уточнить эту формулу.

Решение. Полагаем

$$\tilde{y}=\frac{x}{0{,}051x+0{,}209}+c, \quad \text{где} \quad c=\frac{\sum_{i=1}^{8}\varepsilon_i}{8}=\frac{0{,}47}{8}=0{,}06$$

(см. таблицу 26). Поэтому

$$\tilde{y}=\frac{x}{0{,}051x+0{,}209}+0{,}06=\frac{1{,}0031x+0{,}0125}{0{,}051x+0{,}209}. \tag{7}$$

Результаты уточнения приведены в таблице 26.

Таблица 26

Сравнение эмпирических формул (6) и (7)

x	y	$\bar{y}$ по формуле (6)	$\tilde{y}$ по уточн. ф-ле (7)	$\varepsilon=y-\bar{y}$	ε^2	$\tilde{\varepsilon}=y-\tilde{y}$	$\tilde{\varepsilon}^2$
8	13,0	12,97	13,02	+0,03	$9\cdot10^{-4}$	—0,02	$4\cdot10^{-4}$
10	14,0	13,91	13,96	+0,09	$81\cdot10^{-4}$	+0,04	$16\cdot10^{-4}$
15	15,4	15,40	15,45	0,00	0	—0,05	$25\cdot10^{-4}$
20	16,3	16,27	16,32	+0,03	$9\cdot10^{-4}$	—0,02	$4\cdot10^{-4}$
30	17,2	17,25	17,30	—0,05	$25\cdot10^{-4}$	—0,10	$100\cdot10^{-4}$
40	17,8	17,79	17,84	+0,01	$1\cdot10^{-4}$	—0,04	$16\cdot10^{-4}$
60	18,5	18,39	18,44	+0,11	$121\cdot10^{-4}$	+0,06	$36\cdot10^{-4}$
80	18,8	18,65	18,70	+0,15	$225\cdot10^{-4}$	+0,10	$100\cdot10^{-4}$
Σ				0,47	0,0471	—0,03	0,0301

§ 12. Общий метод определения параметров эмпирической формулы

До сих пор мы рассматривали в основном эмпирические формулы, линейно зависящие от параметров (или приводимые к такому виду), и для этого частного случая давались эффективные методы определения параметров. Сейчас мы укажем общий метод определения параметров эмпирической формулы, не предполагая ее линейности относительно параметров.

Пусть для совокупности значений (x_i, y_i) $(i=1, 2, \ldots, n)$ построена эмпирическая формула

$$y=\tilde{f}(x;\ a_1,\ a_2,\ \ldots,\ a_m), \tag{1}$$

содержащая m параметров ($m < n$), причем функция $\tilde{f}$ имеет непрерывные частные производные по всем своим аргументам. Требуется определить параметры $a_1, a_2, \ldots, a_m$ так, чтобы формула (1) оказалась согласованной с исходными данными, т. е. невязки (уклонения)

$$\varepsilon_i = y_i - \tilde{f}(x_i;\ a_1, a_2, \ldots, a_m) \quad (i = 1, 2, \ldots, n)$$

должны быть возможно малыми по абсолютной величине.

Если бы исходные данные (x_i, y_i) не содержали ошибок и зависимость $\tilde{f}$ была бы точной, то задача отыскания значений параметров $a_i\ (i = 1, 2, \ldots, m)$ свелась бы к решению m уравнений ($m < n$) из следующей системы n уравнений:

$$\left.\begin{aligned} y_1 &= \tilde{f}(x_1;\ a_1, a_2, \ldots, a_m), \\ y_2 &= \tilde{f}(x_2;\ a_1, a_2, \ldots, a_m), \\ &\ldots\ldots\ldots\ldots \\ y_n &= \tilde{f}(x_n;\ a_1, a_2, \ldots, a_m). \end{aligned}\right\} \tag{2}$$

Однако на практике, ввиду отсутствия указанных обстоятельств, система (2) обычно является несовместной, т. е. значения параметров $a_i\ (i = 1, 2, \ldots, m)$, найденные из некоторых m уравнений системы (2), не удовлетворяют остальным $n - m$ уравнениям.

Для приближенного решения системы (2) поступают следующим образом. Каким-либо способом (например, графически или путем решения выбранных m уравнений системы (2)) находят грубые значения параметров $a_1^{(0)}, a_2^{(0)}, \ldots, a_m^{(0)}$.

Пусть

$$a_i = a_i^{(0)} + \alpha_i \quad (i = 1, 2, \ldots, m), \tag{3}$$

где α_i — поправки, которые считаются «малыми», и

$$\varepsilon_j^{(0)} = y_j - \tilde{f}(x_j;\ a_1^{(0)}, a_2^{(0)}, \ldots, a_m^{(0)}) \quad (j = 1, 2, \ldots, n)$$

— соответствующие невязки. Подставляя значения (3) в уравнения системы (2) и разлагая правые части полученных уравнений по степеням поправок α_i, удерживая лишь члены первого порядка относительно этих поправок, будем иметь

$$y_j = \tilde{f}(x_j;\ a_1^{(0)}, a_2^{(0)}, \ldots, a_m^{(0)}) + \sum_{k=1}^{m} \tilde{f}'_{a_k}(x_j;\ a_1^{(0)}, a_2^{(0)}, \ldots, a_m^{(0)})\,\alpha_k$$
$$(j = 1, 2, \ldots, n).$$

Вводя сокращенные обозначения $\tilde{f}'_{a_k}(x_j;\ a_1^{(0)}, a_2^{(0)}, \ldots, a_m^{(0)}) = b_{jk}$ ($j = 1, 2, \ldots, n;\ k = 1, 2, \ldots, m$), находим

$$\sum_{k=1}^{m} b_{jk}\alpha_k = \varepsilon_j^{(0)} \quad (j = 1, 2, \ldots, n). \tag{4}$$

Система (4) линейна относительно неизвестных поправок α_k ($k = 1, 2, \ldots, m$) и является, вообще говоря, несовместной, так как число уравнений ее больше числа неизвестных.

Уравнения системы (4) называются *условными*, а сама система — *системой условных уравнений*.

Система условных уравнений (4) может быть «решена», в известном смысле, описанным выше методом средних (§ 7) или методом наименьших квадратов (§ 8).

Подставляя в нелинейную систему (2) найденные значения

$$a_i^{(1)} = a_i^{(0)} + \alpha_i \quad (i = 1, 2, \ldots, m),$$

можно определить новые невязки $\varepsilon_i^{(1)}$ и в случае необходимости повторить процесс.

Замечание. Если значения y_i в системе (2) получены в результате измерений, то следует позаботиться, чтобы эти значения были равноточными. В противном случае необходимо ввести соответствующие веса [11].

Пример 1. Результаты эксперимента характеризуются следующей таблицей [11]:

x	0	1	2	3
y	2,01	1,21	0,74	0,45

Предполагая, что переменные x и y подчиняются показательной зависимости

$$y = ae^{-bx}, \tag{5}$$

определить наилучшие значения параметров a и b.

Решение. Подставляя табличные данные в формулу (5), получим систему

$$a = 2{,}01, \quad ae^{-b} = 1{,}21, \quad ae^{-b2} = 0{,}74, \quad ae^{-b3} = 0{,}45. \tag{6}$$

Решая, например, первые два уравнения этой системы, будем иметь грубые значения параметров $a^{(0)} = 2{,}01$, $b^{(0)} = 0{,}51$.

Для нахождения поправок $\alpha = a - a^{(0)}$, $\beta = b - b^{(0)}$ составим систему условных уравнений (4). Соответствующие значения производных

$$\left(\frac{\partial y}{\partial a}\right)_0 = e^{-b^{(0)}x}, \quad \left(\frac{\partial y}{\partial b}\right)_0 = -a^{(0)} x e^{-b^{(0)}x}$$

и невязок $\varepsilon^{(0)} = y - a^{(0)} e^{-b^{(0)}x}$ помещены в таблице 27.

Таблица 27

Коэффициенты условных уравнений

x	$\left(\frac{\partial y}{\partial a}\right)_0$	$\left(\frac{\partial y}{\partial b}\right)_0$	$\varepsilon^{(0)}$
0	1	0	0
1	0,600	—1,206	0,004
2	0,361	—1,451	0,014
3	0,216	—1,302	0,016

Отсюда получаем систему условных уравнений

$$\alpha=0,$$
$$\left.\begin{aligned}0{,}600\alpha-1{,}206\beta&=0{,}004,\\0{,}361\alpha-1{,}451\beta&=0{,}014,\\0{,}216\alpha-1{,}302\beta&=0{,}016.\end{aligned}\right\}\qquad(7)$$

Для решения системы (7) применим метод наименьших квадратов. Промежуточные вычисления приведены в таблице 28, где c_α и c_β обозначают коэффициенты при α и β в системе (7) и c_0—соответствующие свободные члены.

Таблица 28

Решение системы (7) методом наименьших квадратов

c_α	c_β	c_0	c_α^2	$c_\alpha c_\beta$	c_β^2	$c_\alpha c_0$	$c_\beta c_0$
1	0	0	1	0	0	0	0
0,600	—1,206	0,004	0,3600	—0,7236	1,4544	0,00240	—0,00482
0,361	—1,451	0,014	0,1303	—0,5238	2,1054	0,00505	—0,02031
0,216	—1,302	0,016	0,0467	—0,2812	1,6952	0,00345	—0,02083
Σ			1,5370	—1,5286	5,2550	0,01090	—0,04596

Следовательно, нормальная система уравнений имеет вид

$$1{,}5370\alpha-1{,}5286\beta=0{,}01090,\qquad -1{,}5286\alpha+5{,}2550\beta=-0{,}04596.$$

Решив эту систему, найдем $\alpha=-0{,}004$, $\beta=-0{,}009$. Отсюда получаем исправленные значения параметров

$$a=2{,}01-0{,}004=2{,}006,\qquad b=0{,}51-0{,}009=0{,}501.$$

Таким образом, искомая эмпирическая формула есть

$$\tilde{y}=2{,}006e^{-0{,}501x}.\qquad(8)$$

В таблице 29 приведено сравнение результатов, даваемых эмпирической формулой (8), с табличными данными.

Таблица 29

Сравнение значений эмпирической формулы (8) с табличными данными

x	y	$\tilde{y}$	$\varepsilon=y-\tilde{y}$	ε^2
0	2,01	2,006	0,004	$1,6\cdot10^{-5}$
1	1,21	1,215	−0,005	$2,5\cdot10^{-5}$
2	0,74	0,736	0,004	$1,6\cdot10^{-5}$
3	0,45	0,446	0,004	$1,6\cdot10^{-5}$

$$\sum \varepsilon_i^2 = 7,3\cdot10^{-5}$$

Заметим, что с помощью логарифмирования формулы (5) легко получить эквивалентную формулу, линейно зависящую от подходящих параметров.

Пример 2. Для переменных x и y, значения которых заданы следующей таблицей:

x	0	0,2	0,6	1,2	1,6	2
y	2,05	1,944	1,638	0,907	0,423	0,028

установлена эмпирическая зависимость

$$\tilde{y}=e^{cx}(a\sin x+b\cos x). \tag{9}$$

Определить значения параметров a, b и c в формуле (9).

Решение. Подставив в эмпирическую формулу (9) данные значения $(x_i,\ y_i)$, получим следующую систему уравнений:

$$\left.\begin{aligned}
2,05 &= b,\\
1,944 &= e^{0,2c}(0,1987a+0,9801b),\\
1,638 &= e^{0,6c}(0,5646a+0,8253b),\\
0,907 &= e^{1,2c}(0,9320a+0,3624b),\\
0,423 &= e^{1,6c}(0,9996a-0,0292b),\\
0,028 &= e^{2c}\ (0,9090a-0,4169b).
\end{aligned}\right\} \tag{10}$$

Решая приближенно, например, первые три уравнения системы (10), найдем $a_0=1$, $b_0=2,05$, $c_0=-0,495$.

Положим $a=a_0+\alpha$, $b=b_0+\beta$, $c=c_0+\gamma$.

Для определения поправок α, β и γ составим линейную систему условных уравнений (4). Значения производных

$$\left(\frac{\partial\tilde{y}}{\partial a}\right)_0=e^{c_0x}\sin x,\qquad \left(\frac{\partial\tilde{y}}{\partial b}\right)_0=e^{c_0x}\cos x,\qquad \left(\frac{\partial\tilde{y}}{\partial c}\right)_0=xe^{c_0x}(a_0\sin x+b_0\cos x)$$

и невязок

$$\varepsilon^{(0)} = y - e^{c_0 x}(a_0 \sin x + b_0 \cos x)$$

помещены в таблице 30.

Таблица 30

Коэффициенты условных уравнений

x	$\left(\frac{\partial \tilde{y}}{\partial a}\right)_0$	$\left(\frac{\partial \tilde{y}}{\partial b}\right)_0$	$\left(\frac{\partial \tilde{y}}{\partial c}\right)_0$	$\varepsilon^{(0)}$
0	0	1	0	0
0,2	0,1795	0,0886	0,3992	—0,0540
0,6	0,4183	0,6114	1,0030	—0,0336
1,2	0,5167	0,2009	1,1141	—0,0214
1,6	0,4536	—0,0132	0,6832	—0,0034
2	0,3378	—0,1549	0,0404	—0,0072

Отсюда система уравнений для поправок имеет вид

$$\beta = 0,$$

$$\left.\begin{aligned} 0{,}1795\alpha + 0{,}0886\beta + 0{,}3992\gamma &= -0{,}0540, \\ 0{,}4183\alpha + 0{,}6114\beta + 1{,}0030\gamma &= -0{,}0336, \\ 0{,}5167\alpha + 0{,}2009\beta + 1{,}1141\gamma &= -0{,}0214, \\ 0{,}4536\alpha - 0{,}0132\beta + 0{,}6823\gamma &= -0{,}0034, \\ 0{,}3378\alpha - 0{,}1549\beta + 0{,}0404\gamma &= -0{,}0072. \end{aligned}\right\} \quad (11)$$

Решение системы (11) будем искать методом средних. Группируя по два уравнения системы, получим

$$\left.\begin{aligned} 0{,}1795\alpha + 1{,}0886\beta + 0{,}3992\gamma &= -0{,}0540, \\ 0{,}9350\alpha + 0{,}8123\beta + 2{,}1171\gamma &= -0{,}0550, \\ 0{,}7914\alpha - 0{,}1681\beta + 0{,}7227\gamma &= -0{,}0106. \end{aligned}\right\} \quad (12)$$

Решая обычным методом систему (12), находим $\alpha = -0{,}0266$; $\beta = -0{,}0466$; $\gamma = 0{,}0036$. Следовательно,

$$a = 1 - 0{,}0266 = 0{,}9734; \quad b = 2{,}05 - 0{,}0466 = 2{,}0034;$$
$$c = -0{,}495 + 0{,}0036 = -0{,}4914.$$

Таким образом, искомая эмпирическая формула имеет вид

$$\tilde{y} = e^{-0{,}4914x}(0{,}9734 \sin x + 2{,}0034 \cos x). \quad (13)$$

В таблице 31 даны расхождения между значениями функции $\tilde{y}$, найденными по эмпирической формуле (13), и табличными значениями y.

Таблица 31

Характеристика точности эмпирической формулы (13)

x	y	$\tilde{y}$	$\Delta = y - \tilde{y}$	x	y	$\tilde{y}$	$\Delta = y - \tilde{y}$
0	2,050	2,003	0,047	1,2	0,907	0,906	0,001
0,2	1,944	1,960	—0,016	1,6	0,423	0,417	0,006
0,6	1,638	1,640	—0,002	2	0,028	0,019	0,007

ЛИТЕРАТУРА К ГЛАВЕ II

[1] Семендяев К. А., Эмпирические формулы, ГТТИ, 1933.
[2] Успенский А. К., Выбор вида и нахождение параметров эмпирической формулы, 1960.
[3] Уорсинг А. и Геффенер Дж., Методы обработки экспериментальных данных, ИЛ, 1949.
[4] Уиттекер Э. и Робинсон Г., Математическая обработка результатов наблюдений, ГТТИ, 1933.
[5] Скарборо Дж., Численные методы математического анализа, ГТТИ, 1934, гл. XVI.
[6] Блох Л. С., Основные графические методы обработки опытных данных, Машгиз, 1951.
[7] Цукерман М. Л., Эмпирические формулы, 1932.
[8] Тумаркин С. А., Об оценке ошибок в методе средних, Труды ЦАГИ, вып. 198 (1935).
[9] Крыжановская М. П., Эмпирические формулы и основы номографии, 1949, § 8.
[10] Фролов С. В., Приближенные вычисления, 1948, Изд. МВТУ, гл. III.
[11] Березин И. С., Жидков Н. П., Методы вычислений, «Наука», 1966, т. I, гл. V.
[12] Батунер Л. М., Позин М. Е., Математические методы в химической технике, Госхимиздат, 1963, гл. XII.
[13] Яковлев К. П., Математическая обработка результатов измерений, Гостехиздат, 1953.
[14] Виньерон А., Обработка результатов физико-химических наблюдений, ОНТИ, 1936.
[15] Демидович Б. П., Марон И. А., Основы вычислительной математики, «Наука», 1966, гл. XIV, § 18.

ГЛАВА III

ПРИБЛИЖЕННОЕ РЕШЕНИЕ ОБЫКНОВЕННЫХ ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ

§ 1. Общие замечания

Простейшим обыкновенным дифференциальным уравнением [1]—[5] является уравнение первого порядка

$$y' = f(x, y). \tag{1}$$

Основная задача, относящаяся к этому уравнению, есть задача Коши: найти решение уравнения (1)

$$y = y(x), \tag{2}$$

удовлетворяющее *начальному условию* $y(x_0) = y_0$; иными словами, требуется найти интегральную кривую $y = y(x)$, проходящую через заданную точку $M_0(x_0, y_0)$ (рис. 21). Если правая часть $f(x, y)$ непрерывна в области R, определяемой неравенствами

$$|x - x_0| < a, \ |y - y_0| < b,$$

то существует по меньшей мере одно решение (2), определенное в *некоторой* окрестности $|x - x_0| < h$, где h — положительное число. Это решение единственно, если в R выполнено *условие Липшица*

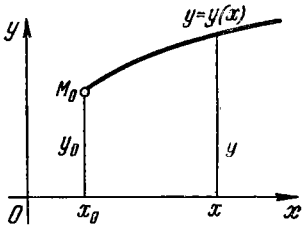

Рис. 21.

$$|f(x, \tilde{y}) - f(x, y)| \leqslant N|\tilde{y} - y|, \tag{3}$$

где N — постоянная (константа Липшица), зависящая в общем случае от a и b. Если $f(x, y)$ имеет ограниченную производную $f'_y(x, y)$ в R, то можно положить

$$N = \max |f'_y(x, y)| \text{ при } (x, y) \in R.$$

Для дифференциального уравнения n-го порядка

$$y^{(n)}=f(x,\ y,\ y',\ \ldots,\ y^{(n-1)})$$

задача Коши состоит в нахождении решения $y=y(x)$, удовлетворяющего *начальным условиям*

$$y(x_0)=y_0,\quad y'(x_0)=y'_0,\ \ldots,\ y^{(n-1)}(x_0)=y_0^{(n-1)},$$

где $x_0,\ y_0,\ y'_0,\ \ldots,\ y_0^{(n-1)}$ — заданные числа.

Пример 1. Дифференциальные уравнения свободных колебаний маятника в сопротивляющейся среде [6], [7].

Пусть Θ — угол отклонения (рис. 22), t — время. Предполагая, что сопротивление среды R пропорционально скорости, получаем для $\Theta=\Theta(t)$ нелинейное дифференциальное уравнение второго порядка

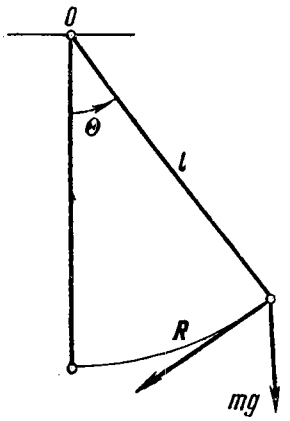

Рис. 22.

$$\frac{d^2\Theta}{dt^2}+a\frac{d\Theta}{dt}+b\sin\Theta=0, \tag{4}$$

где a и b — положительные постоянные.

Начальные условия имеют вид

$$\Theta(t_0)=\Theta_0,\quad \dot\Theta(t_0)=\dot\Theta_0,$$

где

$$\dot\Theta=\frac{d\Theta}{dt},$$

т. е. в начальный момент $t=t_0$ задаются:

1) начальное отклонение Θ_0,

2) начальная угловая скорость $\dot\Theta_0$.

В приложениях часто встречаются системы обыкновенных дифференциальных уравнений. Ограничимся рассмотрением нормальной системы n-го порядка обыкновенных дифференциальных уравнений:

$$\left.\begin{aligned}
\frac{dy_1}{dx}&=f_1(x,\ y_1,\ y_2,\ \ldots,\ y_n),\\
\frac{dy_2}{dx}&=f_2(x,\ y_1,\ y_2,\ \ldots,\ y_n),\\
&\ldots\ldots\ldots\ldots\ldots\\
\frac{dy_n}{dx}&=f_n(x,\ y_1,\ y_2,\ \ldots,\ y_n),
\end{aligned}\right\} \tag{5}$$

где x — независимая переменная, $y_1,\ y_2,\ \ldots,\ y_n$ — искомые функции.

Систему, содержащую производные высших порядков и разрешенную относительно старших производных искомых функций, путем введения новых неизвестных функций можно привести к виду (5). В частности, для дифференциального уравнения n-го порядка

$$y^{(n)}=f(x,\ y,\ y',\ \ldots,\ y^{(n-1)}), \tag{6}$$

полагая $y_1 = y'$, $y_2 = y''$, ..., $y_{n-1} = y^{(n-1)}$, будем иметь эквивалентную нормальную систему

$$\left.\begin{aligned} \frac{dy}{dx} &= y_1, \\ \frac{dy_1}{dx} &= y_2, \\ &\cdots\cdots \\ \frac{dy_{n-2}}{dx} &= y_{n-1}, \\ \frac{dy_{n-1}}{dx} &= f(x, y, y_1, \ldots, y_{n-1}). \end{aligned}\right\} \tag{7}$$

Воспользовавшись векторными обозначениями

$$\boldsymbol{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}, \qquad \frac{d\boldsymbol{y}}{dx} = \begin{bmatrix} \frac{dy_1}{dx} \\ \vdots \\ \frac{dy_n}{dx} \end{bmatrix},$$

систему уравнений (5) можно записать более просто:

$$\frac{d\boldsymbol{y}}{dx} = \boldsymbol{f}(x, \boldsymbol{y}), \tag{8}$$

где $\boldsymbol{f} = \begin{bmatrix} f_1 \\ \vdots \\ f_n \end{bmatrix}$ — заданная вектор-функция.

Под *решением системы* (5) понимается любая совокупность функций

$$\boldsymbol{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}, \tag{9}$$

где $y_1 = \varphi_1(x)$, ..., $y_n = \varphi_n(x)$, которая, будучи подставлена в уравнения (5), обращает их в тождества. Геометрически каждое решение (9) представляет собой некоторую линию L (*интегральная кривая* в пространстве $E^{n+1} = \{x, y_1, \ldots, y_n\}$, рис. 23), а совокупность всех решений образует поле интегральных кривых.

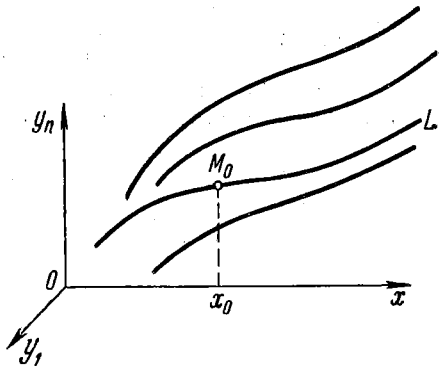

Рис. 23.

Так как система дифференциальных уравнений имеет бесчисленное множество решений, то для выделения одного конкретного

решения $\mathbf{y}=\mathbf{y}(x)$, кроме уравнения, нужны дополнительные условия. В простейшем случае задаются *начальные условия*

$$\mathbf{y}(x_0)=\mathbf{y}^{(0)}, \tag{10}$$

что приводит к *задаче Коши.*

Задача Коши. Найти решение $\mathbf{y}=\mathbf{y}(x)$ системы (5) или соответствующего векторного уравнения (8), удовлетворяющее заданным начальным условиям (10), где x_0 — фиксированное значение независимой переменной и

$$\mathbf{y}^{(0)}=\begin{bmatrix} y_1^{(0)} \\ \vdots \\ y_n^{(0)} \end{bmatrix}$$

— данная система чисел.

Геометрически это значит, что требуется отыскать интегральную кривую L, проходящую через заданную точку $M_0(x_0, y_1^{(0)}, \ldots, y_n^{(0)})$ пространства E^{n+1} (рис. 23).

Если x интерпретировать как время, а $y_1, \ldots, y_n$ — как обобщенные координаты некоторой механической системы, то получим следующий аспект задачи Коши: зная дифференциальные уравнения, управляющие механической системой, а также состояние ее в начальный момент времени x_0, определить состояние системы в любой момент времени x.

Гарантия однозначной разрешимости задачи Коши дается приведенными ниже достаточными условиями.

Теорема существования и единственности решения. *Пусть в некоторой окрестности начальных значений*

$$U\{|x-x_0|<a, \quad |y_1-y_1^{(0)}|<b_1, \ldots, |y_n-y_n^{(0)}|<b_n\}$$

система (5) *обладает следующими свойствами:*

1) *правые части* $f_1, f_2, \ldots, f_n$ *определены и непрерывны в* U;

2) *функции* f_i $(i=1, 2, \ldots, n)$ *в окрестности* U *удовлетворяют условиям Липшица по зависимым переменным* $y_1, y_2, \ldots, y_n$, *т. е.*

$$|f_i(x, \tilde{y}_1, \ldots, \tilde{y}_n)-f_i(x, y_1, \ldots, y_n)|\leqslant N\sum_{j=1}^{n}|\tilde{y}_j-y_j| \tag{11}$$
$$(i=1, 2, \ldots, n),$$

где

$$(x, y_1, \ldots, y_n)\in U, \quad (x, \tilde{y}_1, \ldots, \tilde{y}_n)\in U$$

и N — *некоторая постоянная (константа Липшица). В этом случае существует единственное решение системы* (5) $y_1=y_1(x), \ldots$

..., $y_n = y_n(x)$, определенное в некоторой окрестности $|x - x_0| < h$ и удовлетворяющее заданным начальным условиям:

$$y_1(x_0) = y_1^{(0)}, \ \ldots, \quad y_n(x_0) = y_n^{(0)}.$$

Иными словами, при выполнении условий 1), 2) задача Коши разрешима и решение единственно, т. е. через точку $M_0(x_0, y_1^{(0)}, \ldots, y_n^{(0)})$ проходит единственная интегральная кривая системы (5).

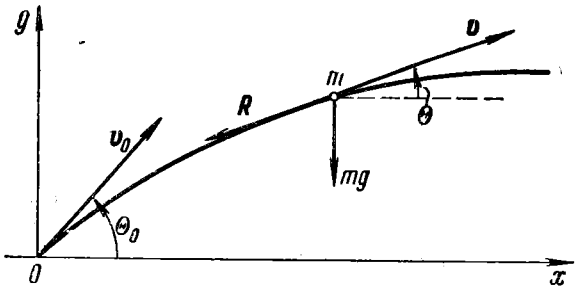

Рис. 24.

Заметим, что вместо условий Липшица достаточно потребовать наличия ограниченных производных $\frac{\partial f_i}{\partial y_j}$ $(i, j = 1, 2, \ldots, n)$ в окрестности U; тогда за константу Липшица можно принять

$$N = \max_{i,\, j} \left| \frac{\partial f_i}{\partial y_j} \right|.$$

Рассмотрим примеры некоторых систем дифференциальных уравнений.

Пример 2. (Основная задача внешней баллистики [8].) Приведем дифференциальные уравнения движения материальной точки (снаряда) в сопротивляющейся среде (рис. 24).

Пусть m — масса точки; x, y — текущие координаты точки в вертикальной плоскости Oxy; t — время; $\boldsymbol{v}$ — вектор скорости; Θ — угол, образованный вектором скорости с горизонтальной осью Ox; $\boldsymbol{R}$ — вектор сопротивления среды (направлен по касательной к траектории); g — ускорение силы тяжести.

На основании второго закона Ньютона получаем дифференциальные уравнения движения

$$m\frac{d^2x}{dt^2} = -R\cos\Theta, \quad m\frac{d^2y}{dt^2} = -mg - R\sin\Theta. \tag{12}$$

Обычно предполагают [8], что величина силы сопротивления среды

$$R = mF(v)\,H(y), \tag{13}$$

где множитель $F(v)$ зависит только от величины скорости $v=|\boldsymbol{v}|$, а второй $H(y)$ — от высоты поднятия y. Функции $F(v)$ и $H(y)$ задаются таблично, с учетом реальной обстановки.

Подставляя выражение (13) в систему (12), после сокращения на m будем иметь

$$\left.\begin{aligned}\frac{d^2x}{dt^2}&=-H(y)\,F(v)\cos\Theta,\\ \frac{d^2y}{dt^2}&=-g-H(y)\,F(v)\sin\Theta.\end{aligned}\right\}\qquad(14)$$

Чтобы привести систему (14) к нормальному виду (5), введем новые переменные

$$u=v\cos\Theta,\quad w=v\sin\Theta,$$

которые представляют собой соответственно горизонтальную ($u=v_x$) и вертикальную ($w=v_y$) проекции скорости $\boldsymbol{v}$.

Вводя обозначение $G(v)=\dfrac{F(v)}{v}$ и учитывая, что $u=\dfrac{dx}{dt}$, $w=\dfrac{dy}{dt}$, вместо системы (14) получим нормальную систему четвертого порядка:

$$\left.\begin{aligned}\frac{dx}{dt}&=u,\\ \frac{dy}{dt}&=w,\\ \frac{du}{dt}&=-H(y)\,G(v)\,u,\\ \frac{dw}{dt}&=-g-H(y)\,G(v)\,w,\end{aligned}\right\}\qquad(15)$$

где $v=\sqrt{u^2+w^2}$.

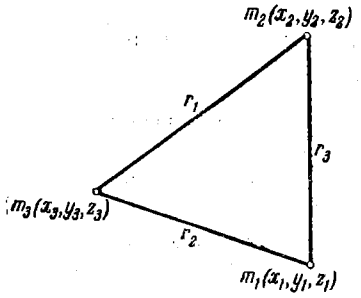

Рис. 25.

Если предположить, что снаряд был брошен в момент времени $t=0$ из точки $O(0, 0)$ со скоростью, по величине равной v_0 и направленной под углом Θ_0 к горизонту, то начальные условия будут иметь вид

$$x(0)=0,\quad y(0)=0,\quad u(0)=v_0\cos\Theta_0,\quad w(0)=v_0\sin\Theta_0.\qquad(16)$$

Считая, что функции $H(y)$ и $G(v)$ непрерывно дифференцируемы, легко убедиться, что задача Коши (15), (16) имеет единственное решение, т. е. начальные условия (16) однозначно определяют траекторию снаряда.

В общем случае система (15) элементарно не интегрируется, тем более что функции $H(y)$ и $G(v)$ носят эмпирический характер и задаются таблично. Поэтому систему (15) приходится решать численно.

Обычно интересуются еще максимальной высотой подъема снаряда $y_{\max}=y(t_1)$ и дальностью полета $x_{\max}=x(t_2)$. Для нахождения этих величин нужно соответственно решить *краевые задачи* (см. гл. IV)

$$y'(t_1)=0\quad\text{и}\quad y(t_2)=0.$$

Пример 3. (Задача трех тел [9]). Речь идет о дифференциальных уравнениях движения трех материальных тел, взаимодействующих между собой по закону всемирного тяготения Ньютона.

Пусть в пространстве $Oxyz$ имеются три тела с массами m_1, m_2 и m_3 (рис. 25), занимающие в момент времени t соответственно положения x_i, y_i, z_i ($i=1, 2, 3$). Конкретной моделью нашей задачи может служить, например, система Земля—Луна—спутник.

В механике доказывается [9], что для системы m_i ($i=1, 2, 3$) существует потенциальная функция

$$U=\frac{m_1m_2}{r_3}+\frac{m_2m_3}{r_1}+\frac{m_3m_1}{r_2}$$

Отсюда дифференциальные уравнения движения системы трех тел будут

$$m_i\frac{d^2x_i}{dt^2}=\frac{\partial U}{\partial x_i},\quad m_i\frac{d^2y_i}{dt^2}=\frac{\partial U}{\partial y_i},\quad m_i\frac{d^2z_i}{dt^2}=\frac{\partial U}{\partial z_i}\quad (i=1,\ 2,\ 3). \tag{17}$$

Полагая

$$\frac{dx_i}{dt}=u_i,\quad \frac{dy_i}{dt}=v_i,\quad \frac{dz_i}{dt}=w_i,$$

будем иметь нормальную систему 18-го порядка:

$$\left.\begin{aligned} &\frac{dx_i}{dt}=u_i, && \frac{dy_i}{dt}=v_i, && \frac{dz_i}{dt}=w_i,\\ &m_i\frac{du_i}{dt}=\frac{\partial U}{\partial x_i}, && m_i\frac{dv_i}{dt}=\frac{\partial U}{\partial y_i}, && m_i\frac{dw_i}{dt}=\frac{\partial U}{\partial z_i}\end{aligned}\right\}\quad (i=1,\ 2,\ 3). \tag{18}$$

Начальные условия сводятся к заданию системы восемнадцати чисел

$$\left.\begin{aligned} &x_i(t_0)=x_i^{(0)}, && y_i(t_0)=y_i^{(0)}, && z_i(t_0)=z_i^{(0)},\\ &u_i(t_0)=u_i^{(0)}, && v_i(t_0)=v_i^{(0)}, && w_i(t_0)=w_i^{(0)}\end{aligned}\right\}(i=1,\ 2,\ 3) \tag{19}$$

— начальных положений тел и проекций их начальных скоростей.

Решив задачу Коши (18), (19), получим законы движения тел в конечной форме:

$$x_i=\varphi_i(t),\quad y_i=\psi_i(t),\quad z_i=\chi_i(t)\qquad (i=1,\ 2,\ 3)$$

и дополнительно законы изменения их скоростей:

$$u_i=\varphi_i'(t),\quad v_i=\psi_i'(t),\quad w_i=\chi_i'(t)\qquad (i=1,\ 2,\ 3).$$

Даже для простейшего дифференциального уравнения первого порядка (1) нахождение решения, отвечающего заданным условиям, вообще невыполнимо с помощью конечного числа математических операций. Тем более это неосуществимо для системы дифференциальных уравнений.

Указанное обстоятельство привело к созданию большого числа методов *приближенного решения* дифференциальных уравнений, основанных на самых различных идеях.

Все эти методы в зависимости от формы, в которой они представляют решение, в основном можно разделить на три группы:

1. Аналитические методы, дающие приближенное решение дифференциального уравнения в виде аналитического выражения.

2. Графические методы, дающие приближенное решение в виде графика.

3. Численные методы, дающие приближенное решение в виде таблицы.

Следует отметить, что приведенная классификация методов приближенного интегрирования является в известной мере условной. Так, например, *графический метод ломаных Эйлера* дает одновременно способ численного решения дифференциального уравнения.

При дальнейшем изложении будем предполагать, не оговаривая этого каждый раз, что для рассматриваемых дифференциальных уравнений выполнены обычные условия существования и единственности решений. Для применимости некоторых методов потребуются более жесткие условия, которые будут указаны в соответствующих местах.

§ 2. Интегрирование дифференциальных уравнений с помощью степенных рядов

Рассмотрим сначала дифференциальное уравнение первого порядка

$$y' = f(x, y) \tag{1}$$

с начальным условием

$$y(x_0) = y_0. \tag{2}$$

Пусть правая часть уравнения (1) является аналитической функцией в начальной точке (x_0, y_0), т. е. в некоторой окрестности этой точки может быть разложена в степенной ряд вида

$$f(x, y) = \sum_{p,\, q=0}^{\infty} c_{pq} (x - x_0)^p (y - y_0)^q,$$

где p, q — целые неотрицательные числа и c_{pq} — постоянные коэффициенты. Тогда существует единственное решение $y = y(x)$ дифференциального уравнения (1), удовлетворяющее начальному условию (2), причем это решение является аналитическим в точке x_0 и, следовательно, может быть представлено в виде ряда Тейлора [2]

$$y(x) = \sum_{p=0}^{\infty} c_p (x - x_0)^p \qquad (|x - x_0| < h), \tag{3}$$

где

$$c_p = \frac{1}{p!} y^{(p)}(x_0) \qquad (p = 0, 1, 2, \ldots)$$

и h — некоторое положительное число.

Коэффициент c_0 разложения (3) определяется непосредственно из начального условия (2):

$$c_0 = y(x_0) = y_0;$$

следующий коэффициент c_1 находится на основании дифференциального уравнения (1):

$$c_1 = y'(x_0) = f(x_0, y_0).$$

Что касается остальных коэффициентов c_p $(p > 1)$ ряда (3), то они могут быть шаг за шагом найдены путем последовательного дифференцирования данного дифференциального уравнения (1). Например, дифференцируя по x обе части уравнения (1) и используя правило дифференцирования сложной функции, будем иметь

$$y'' = f'_x(x, y) + f'_y(x, y)\, y'.$$

Отсюда

$$c_2 = \frac{1}{2} y''(x_0) = \frac{1}{2}\left[f'_x(x_0, y_0) + f'_y(x_0, y_0)\, y'_0\right],$$

где число $y'_0 = f(x_0, y_0)$ уже известно.

Далее находим

$$c_3 = \frac{1}{6} y'''(x_0) =$$

$$= \frac{1}{6}\left[f''_{xx}(x_0, y_0) + 2f''_{xy}(x_0, y_0)\, y'_0 + f''_{yy}(x_0, y_0)\, y'^2_0 + f'_y(x_0, y_0)\, y''_0\right].$$

Аналогично могут быть определены коэффициенты c_4, c_5 и т. д. и, следовательно, формально построено аналитическое решение $y(x)$.

Вопрос об оценке радиуса сходимости h степенного ряда (3) более сложен [2], [11], и здесь мы его рассматривать не будем. Заметим только, что если дифференциальное уравнение линейное $y' = f_0(x) + f_1(x)\, y$ и функции $f_0(x)$, $f_1(x)$ допускают разложения в степенные ряды в начальной точке x_0 с общим радиусом сходимости R, то можно положить $h = R$.

Пример 1. Написать несколько членов разложения в степенной ряд решения $y = y(x)$ дифференциального уравнения

$$y' = \frac{y}{x+y}, \tag{4}$$

удовлетворяющего начальному условию: $y(1) = 2$.

Пользуясь полученным разложением, приближенно найти $y(1, 2)$.

Решение. Полагая $x_0 = 1$ и

$$y = y_0 + y'_0(x - x_0) + \frac{1}{2!} y''_0 (x - x_0)^2 + \frac{1}{3!} y'''_0 (x - x_0)^3 + \ldots$$

$$(y^{(p)}_0 = y^{(p)}(x_0);\ p = 0, 1, 2, \ldots),$$

будем иметь

$$y_0 = 2, \quad y'_0 = \frac{2}{1+2} = \frac{2}{3}.$$

Дифференцируя уравнение (4), получим

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

Отсюда

$$y_0''=\frac{1\cdot\frac{2}{3}-2}{(1+2)^2}=-\frac{4}{27}.$$

Дифференцируя еще раз, будем иметь

$$y'''=\frac{xy''(x+y)-2(1+y')(xy'-y)}{(x+y)^3}.$$

Поэтому

$$y_0'''=\frac{-\frac{4}{27}\cdot 3+2\cdot\frac{5}{3}\cdot\frac{4}{3}}{27}=\frac{4}{27}.$$

Аналогично могут быть найдены остальные производные $y_0^{(p)}$ $(p>3)$.

Таким образом,

$$y=2+\frac{2}{3}(x-1)-\frac{2}{27}(x-1)^2+\frac{2}{81}(x-1)^3+\ldots \tag{5}$$

Из формулы (5) имеем $y(1,\ 2)=2+0{,}1333-0{,}0030+0{,}0002+\ldots\approx 2{,}1305$.

Изложенный выше способ нахождения решения дифференциального уравнения методом степенных рядов легко обобщается на случай дифференциального уравнения n-го порядка. Пусть, например, мы имеем дифференциальное уравнение второго порядка

$$y''=f(x,\ y,\ y') \tag{6}$$

с начальными условиями:

$$y(x_0)=y_0,\quad y'(x_0)=y_0'. \tag{7}$$

Предполагая, что функция $f(x,\ y,\ y')$ аналитическая в начальной точке $(x_0,\ y_0,\ y_0')$, будем искать решение задачи Коши (6), (7) в виде ряда Тейлора

$$y(x)=\sum_{p=0}^{\infty}\frac{y_0^{(p)}}{p!}(x-x_0)^p. \tag{8}$$

Здесь y_0 и y_0' известны из начальных условий. Из уравнения (6) получаем

$$y_0''=f(x_0,\ y_0,\ y_0').$$

Дифференцируя последовательно уравнение (6) по переменной x согласно правилу дифференцирования сложной функции и полагая $x=x_0$, будем иметь

$$y_0'''=f_x'(x_0,\ y_0,\ y_0')+f_y'(x_0,\ y_0,\ y_0')\,y_0'+f_{y'}'(x_0,\ y_0,\ y_0')\,y_0''$$

и т. д. Таким образом шаг за шагом может быть построен ряд (8).

Вопрос о сходимости этого ряда оставим открытым.

Пример 2. Написать несколько членов разложения в степенной ряд решения $y=y(x)$ уравнения

$$y''+xy'+y=0, \tag{9}$$

удовлетворяющего начальным условиям: $y(0)=0$, $y'(0)=1$.

Решение. Полагаем

$$y(x)=y(0)+y'(0)\,x+\frac{y''(0)}{2!}\,x^2+\ldots,$$

где $y(0)=0$ и $y'(0)=1$.

Из уравнения (9) получаем

$$y''=-xy'-y. \tag{10}$$

Отсюда $y''(0)=-y(0)=0$. Дифференцируя последовательно уравнение (10), будем иметь

$$y'''=-xy''-2y',\quad y^{\mathrm{IV}}=-xy'''-3y'',\quad y^{\mathrm{V}}=-xy^{\mathrm{IV}}-4y''',\ \ldots$$

Из этих равенств вытекает, что

$$y'''(0)=-2\cdot 1=-2,\quad y^{\mathrm{IV}}(0)=-3\cdot 0=0,\quad y^{\mathrm{V}}(0)=-4\cdot(-2)=8.$$

Следовательно,

$$y(x)=x-\frac{x^3}{3}+\frac{x^5}{15}+\ldots \tag{11}$$

Написать общий член ряда (11) и исследовать его сходимость не представляет больших затруднений.

Пример 3. Зная дифференциальное уравнение движения точки [10]

$$\frac{d^2x}{dt^2}+(1+0{,}1t)\,x+0{,}1\left(\frac{dx}{dt}\right)^2=0, \tag{12}$$

с помощью разложения решения в степенной ряд найти скорость x' и ускорение x'' точки для моментов времени t, равных 0,1; 0,2; 0,3; 0,4; 0,5, если $x(0)=1$, $x'(0)=2$.

Решение. Из уравнения (12) будем иметь

$$x''=-x-0{,}1tx-0{,}1x'^2. \tag{13}$$

Отсюда с помощью последовательного дифференцирования получим

$$\begin{aligned}
x''' &= -x'-0{,}1\,(tx'+x)-0{,}2x'x'',\\
x^{\mathrm{IV}} &= -x''-0{,}1\,(tx''+2x')-0{,}2\,(x'x'''+x''^2),\\
x^{\mathrm{V}} &= -x'''-0{,}1\,(tx'''+3x'')-0{,}2\,(x'x^{\mathrm{IV}}+3x''x'''),\\
x^{\mathrm{VI}} &= -x^{\mathrm{IV}}-0{,}1\,(tx^{\mathrm{IV}}+4x''')-0{,}2\,(x'x^{\mathrm{IV}}+4x''x^{\mathrm{IV}}+3x'''^2)
\end{aligned}$$

и т. д.

Полагая в этих равенствах $t=0$ и используя начальные условия $x_0=1$, $x_0'=2$, находим

$$x_0''=-1{,}4,\quad x_0'''=-1{,}54,\quad x_0^{\mathrm{IV}}=1{,}224,\quad x_0^{\mathrm{V}}=0{,}1768,\quad x_0^{\mathrm{VI}}=-0{,}7308.$$

Так как на основании ряда Тейлора

$$x=x_0+x_0't+\frac{x_0''t^2}{2!}+\frac{x_0'''t^3}{3!}+\frac{x_0^{\mathrm{IV}}t^4}{4!}+\frac{x_0^{\mathrm{V}}t^5}{5!}+\frac{x_0^{\mathrm{VI}}t^6}{6!}+\ldots,$$

то для искомого решения с точностью до t^6 получаем

$$x = 1 + 2t - 0{,}7t^2 - 0{,}2567t^3 + 0{,}051t^4 + 0{,}00147t^5 - 0{,}00101t^6. \qquad (14)$$

Следовательно,

$$x' = 2 - 1{,}4t - 0{,}77t^2 + 0{,}204t^3 + 0{,}00735t^4 - 0{,}00606t^5, \qquad (15)$$

$$x'' = -1{,}4 - 1{,}54t + 0{,}612t^2 + 0{,}00294t^3 - 0{,}0303t^4. \qquad (16)$$

В таблице 32 помещены значения x, x' и x'' для t, равного 0; 0,1; 0,2; 0,3; 0,4; 0,5, подсчитанные соответственно по формулам (14), (15) и (16). Для контроля приведены значения $\tilde{x}''$, полученные непосредственно из точного уравнения (13).

Таблица 32

Вычисление значений решения x дифференциального уравнения (12) и его производных x' и x'' при помощи разложения в степенной ряд

t	x	x'	x''	$\tilde{x}''$	$x''-\tilde{x}''$
0	1,000	2,000	−1,400	−1,400	0,000
0,1	1,193	1,852	−1,548	−1,549	+0,001
0,2	1,370	1,692	−1,683	−1,684	+0,001
0,3	1,530	1,517	−1,806	−1,806	0,000
0,4	1,673	1,330	−1,917	−1,917	0,000
0,5	1,796	1,137	−2,015	−2,015	0,000

Таким образом, видно, что при $0 \leqslant t \leqslant 1/2$ формулы (14), (15) и (16) дают решение задачи, точное до третьего десятичного знака. При увеличении t точность этих формул будет, вообще говоря, убывать, и при больших t придется учитывать дополнительные члены ряда Тейлора.

Заметим, что формула (14) совершенно непригодна для анализа устойчивости движения при $t \to \infty$; для этой цели используют ряды иного вида.

Метод степенных рядов применим также к нормальным системам дифференциальных уравнений

$$\frac{dy_i}{dx} = f_i(x, y_1, \ldots, y_n) \qquad (i = 1, \ldots, n)$$

с начальными условиями $y_i(x_0) = y_{i0}$ $(i = 1, \ldots, n)$.

Если функции $f_i(x, y_1, \ldots, y_n)$ $(i = 1, \ldots, n)$ аналитические, то компоненты решения $y = \{y_1(x), \ldots, y_n(x)\}$ системы можно искать в виде рядов Тейлора

$$y_i(x) = \sum_{p=0}^{\infty} \frac{y_{i0}^{(p)}(x_0)}{p!}(x - x_0)^p \qquad (i = 1, \ldots, n).$$

Коэффициенты этих рядов могут быть найдены с учетом начальных условий последовательным дифференцированием уравнений данной системы.

Пример 4. Для системы

$$\frac{dx}{dt}=x\cos t-y\sin t,\qquad \frac{dy}{dt}=x\sin t+y\cos t \tag{17}$$

построить решение в форме степенного ряда, удовлетворяющее начальным условиям $x(0)=1$, $y(0)=0$.

Решение. Положим

$$x(t)=x(0)+x'(0)\,t+\frac{x''(0)}{2!}\,t^2+\frac{x'''(0)}{3!}\,t^3+\ldots \tag{18}$$

и

$$y(t)=y(0)+y'(0)\,t+\frac{y''(0)}{2!}\,t^2+\frac{y'''(0)}{3!}\,t^3+\ldots \tag{19}$$

Из начальных условий имеем $x(0)=1$, $y(0)=0$. Полагая $t=0$ в системе (17), получим $x'(0)=1$, $y'(0)=0$. Дифференцируя по t систему (17), будем иметь

$$\left.\begin{aligned}\frac{d^2x}{dt^2}&=-x\sin t-y\cos t+\frac{dx}{dt}\cos t-\frac{dy}{dt}\sin t,\\ \frac{d^2y}{dt^2}&=x\cos t-y\sin t+\frac{dx}{dt}\sin t+\frac{dy}{dt}\cos t.\end{aligned}\right\} \tag{20}$$

Отсюда $x''(0)=1$, $y''(0)=1$. Дифференцируя систему (20), получим

$$\frac{d^3x}{dt^3}=-x\cos t+y\sin t+2\left(-\frac{dx}{dt}\sin t-\frac{dy}{dt}\cos t\right)+\frac{d^2x}{dt^2}\cos t-\frac{d^2y}{dt^2}\sin t,$$

$$\frac{d^3y}{dt^3}=-x\sin t-y\cos t+2\left(\frac{dx}{dt}\cos t-\frac{dy}{dt}\sin t\right)+\frac{d^2x}{dt^2}\sin t+\frac{d^2y}{dt^2}\cos t.$$

Следовательно, $x'''(0)=-1+1=0$, $y'''(0)=3$. Аналогичным путем могут быть найдены и дальнейшие производные.

Используя формулы (18) и (19), окончательно имеем

$$x(t)=1+t+\frac{1}{2}\,t^2+\ldots,\qquad y(t)=\frac{1}{2}\,t^2+\frac{1}{2}\,t^3+\ldots \tag{21}$$

Из формул (21) можно в окрестности начальной точки $t=0$ приближенно найти численные значения искомого решения. Например:

$$x(0{,}1)=1+0{,}1+\frac{1}{2}\cdot 0{,}01=1{,}105,$$

$$y(0{,}1)=\frac{1}{2}\cdot 0{,}01+\frac{1}{2}\cdot 0{,}001=0{,}0055 \text{ и т. д.}$$

Метод разложения решения дифференциального уравнения в степенные ряды часто используется как элемент более практичных методов приближенного интегрирования дифференциальных уравнений. В частности, для некоторых численных методов интегрирования

дифференциальных уравнений требуется определить значения искомых функций в нескольких точках. Эти значения при соблюдении известных условий гладкости данного уравнения могут быть с любой степенью точности подсчитаны с помощью степенных рядов.

§ 3. Метод последовательных приближений

Сначала изложим этот метод применительно к дифференциальному уравнению первого порядка

$$y' = f(x, y) \tag{1}$$

с начальным условием

$$y(x_0) = y_0. \tag{2}$$

Предполагается, что в некоторой окрестности точки $M_0(x_0, y_0)$ уравнение (1) удовлетворяет условиям теоремы существования и единственности решения.

Будем строить искомое решение $y = y(x)$ для значений $x \geqslant x_0$. Случай $x \leqslant x_0$ вполне аналогичен. Интегрируя правую и левую части уравнения (1) в пределах от x_0 до x, получим

$$y(x) - y(x_0) = \int_{x_0}^{x} f(x, y)\, dx$$

или, в силу начального условия (2), будем иметь

$$y(x) = y_0 + \int_{x_0}^{x} f(x, y)\, dx. \tag{3}$$

Так как искомая функция $y = y(x)$ находится под знаком интеграла, то уравнение (3) является интегральным. Очевидно, решение интегрального уравнения (3) удовлетворяет дифференциальному уравнению (1) и начальному условию (2).

Для нахождения этого решения применим метод последовательных приближений. Заменяя в равенстве (3) неизвестную функцию y данным значением y_0, получим первое приближение

$$y_1 = y_0 + \int_{x_0}^{x} f(x, y_0)\, dx.$$

Далее, подставив в равенство (3) вместо неизвестной функции y найденную функцию y_1, будем иметь второе приближение

$$y_2 = y_0 + \int_{x_0}^{x} f(x, y_1)\, dx.$$

Вообще, все дальнейшие приближения строятся по формуле

$$y_n = y_0 + \int_{x_0}^{x} f(x, y_{n-1})\,dx \quad (n = 1, 2, \ldots). \tag{4}$$

Геометрически последовательные приближения представляют собой кривые $y_n = y_n(x)$ $(n = 1, 2, \ldots)$, проходящие через общую точку $M_0(x_0, y_0)$ (рис. 26).

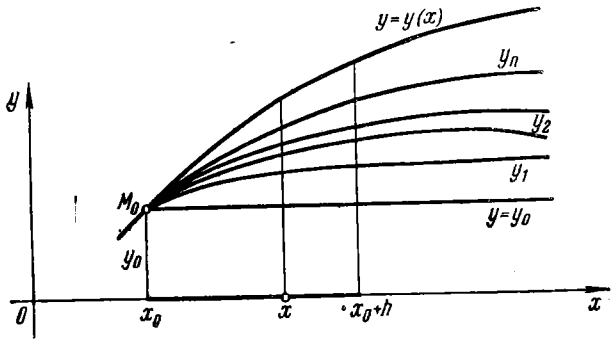

Рис. 26.

В учебниках по дифференциальным уравнениям [1], [2], [4] доказывается, что при выполнении условия Липшица

$$|f(x, \tilde{y}) - f(x, y)| \leqslant N|\tilde{y} - y| \tag{5}$$

последовательные приближения $y_n = y_n(x)$ на некотором достаточно малом отрезке $[x_0, x_0 + h]$ имеют смысл и равномерно сходятся, причем предельная функция

$$y(x) = \lim_{n \to \infty} y_n(x) \tag{6}$$

удовлетворяет дифференциальному уравнению (1) и начальному условию (2).

Если правая часть $f(x, y)$ дифференциального уравнения (1) определена и непрерывна в области $R\{0 \leqslant x - x_0 \leqslant a,\ |y - y_0| \leqslant b\}$ и $M \geqslant \max|f(x, y)|$ при $(x, y) \in R$, то за величину h можно принять [1]

$$h = \min\left(a, \frac{b}{M}\right), \tag{7}$$

причем интегральная кривая $y = y(x)$ при $x_0 \leqslant x \leqslant x_0 + h$ будет

содержаться в угле между прямыми (рис. 27)

$$y = y_0 + M(x - x_0) \text{ и } y = y_0 - M(x - x_0).$$

Для оценки погрешности

$$\varepsilon_n(x) = |y(x) - y_n(x)|$$

из формулы (3) вычтем формулу (4), тогда будем иметь

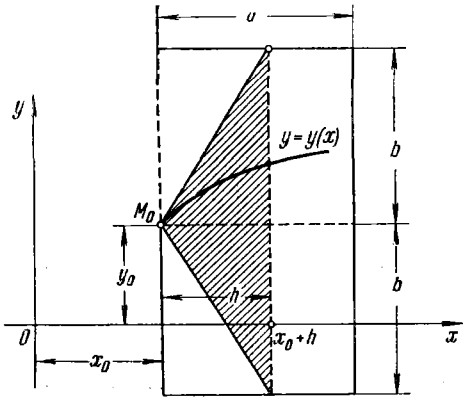

Рис. 27.

$$y(x) - y_n(x) = \int_{x_0}^{x} [f(x, y) - f(x, y_{n-1})].$$

Отсюда при $x_0 \leqslant x \leqslant x_0 + h$ получим

$$\varepsilon_n(x) = |y(x) - y_n(x)| \leqslant \int_{x_0}^{x} |f(x, y) - f(x, y_{n-1})| dx.$$

В силу условия Липшица (5) находим

$$|f(x, y) - f(x, y_{n-1})| \leqslant N|y(x) - y_{n-1}(x)| \leqslant N\varepsilon_{n-1}(x).$$

Следовательно,

$$\varepsilon_n(x) \leqslant \int_{x_0}^{x} N\varepsilon_{n-1}(x)\, dx \quad (n = 1, 2, \ldots), \tag{8}$$

где введено обозначение $\varepsilon_0(x) = |y(x) - y_0|$. Применяя формулу Лагранжа при $x_0 \leqslant x \leqslant x_0 + h$, будем иметь

$$\varepsilon_0(x) = |y(x) - y(x_0)| = (x - x_0)|y'(\xi)|,$$

где $x_0 < \xi < x$.

Отсюда, так как $|y'(\xi)| = |f(\xi, y(\xi))| \leqslant M$, получим

$$\varepsilon_0(x) \leqslant M(x - x_0).$$

Далее, используя формулу (8), последовательно находим

$$\varepsilon_1(x) \leqslant N\int_{x_0}^{x} \varepsilon_0(x)\, dx \leqslant NM\int_{x_0}^{x} (x - x_0)\, dx = MN\frac{(x - x_0)^2}{2!},$$

$$\varepsilon_2(x) \leqslant N\int_{x_0}^{x} \varepsilon_1(x)\, dx \leqslant MN^2\int_{x_0}^{x} \frac{(x - x_0)^2}{2!}\, dx = MN^2\frac{(x - x_0)^3}{3!}$$

и т. д. В итоге окончательно получим

$$\varepsilon_n(x) \leqslant MN^n \frac{(x-x_0)^{n+1}}{(n+1)!} \quad (n=0,\ 1,\ 2,\ \ldots). \tag{9}$$

Из формулы (9) вытекает, что $\varepsilon_n(x) \longrightarrow 0$ при $n \longrightarrow \infty$ равномерно на отрезке $[x_0,\ x_0+h]$.

Замечание. При методе последовательных приближений в качестве начального приближения y_0, вообще говоря, можно выбирать любую функцию, достаточно близкую к точному решению y.

Например, иногда выгодно в качестве y_0 брать конечный отрезок ряда Тейлора искомого решения.

Заметим, что при пользовании методом последовательных приближений аналитичность правой части дифференциального уравнения не обязательна, поэтому метод этот можно применять и в тех случаях, когда разложение решения дифференциального уравнения в степенной ряд невозможно.

Пример 1. Методом последовательных приближений найти приближенное решение дифференциального уравнения

$$y' = x - y, \tag{10}$$

удовлетворяющее начальному условию $y(0)=1$.

Решение. В качестве начального приближения возьмем $y_0(x)=1$. Так как

$$y = 1 + \int_0^x (x-y)\,dx,$$

то будем иметь

$$y_1 = 1 + \int_0^x (x-1)\,dx = 1 - x + \frac{x^2}{2}.$$

Аналогично

$$y_2 = 1 + \int_0^x \left(x - 1 + x - \frac{x^2}{2}\right) dx = 1 - x + x^2 - \frac{x^3}{6}.$$

Подобным же образом получим

$$y_3 = 1 - x + x^2 - \frac{x^3}{3} + \frac{x^4}{24},$$

$$y_4 = 1 - x + x^2 - \frac{x^3}{3} + \frac{x^4}{12} - \frac{x^5}{120}$$

и т. д.

Оценим погрешность, например, четвертого приближения $y_4 = y_4(x)$. Рассмотрим некоторую область $R\{0 \leqslant x \leqslant a,\ |y| \leqslant b\}$, где правая часть дифференциального уравнения (10)

$$f(x,\ y) = x - y \tag{11}$$

определена и непрерывна.

Так как функция (11) непрерывна во всей плоскости Oxy, то за a и b могут быть взяты любые положительные числа.

При $(x, y) \in R$ имеем

$$|f(x, y)| \leqslant |x-y| \leqslant |x|+|y| \leqslant a+b=M.$$

Поэтому, предполагая $a \geqslant 1$, из формулы (7) получаем

$$h=\min\left(a, \frac{b}{M}\right)=\min\left(a, \frac{b}{a+b}\right)=\frac{b}{a+b}.$$

Выбрав для определенности $a=1$ и $b=1$, будем иметь $h=1/2$.

Константа Липшица для области R в данном случае будет

$$N=\max|f'_y(x, y)|=1.$$

Используя формулу (9) при $0 \leqslant x \leqslant 1/2$, окончательно получим

$$\varepsilon_4(x)=|y(x)-y_4(x)| \leqslant 2 \cdot 1^4 \cdot \frac{x^5}{5!}=\frac{x^5}{60}$$

и, следовательно, $\varepsilon_4=\max \varepsilon_4(x) \leqslant \frac{1}{60 \cdot 32}=\frac{1}{1920}$. Нетрудно видеть, что

$$y_n=1-x+2\left[\frac{x^2}{2!}-\frac{x^3}{3!}+\ldots+(-1)^n \frac{x^n}{n!}\right]+(-1)^{n+1} \frac{x^{n+1}}{(n+1)!}$$

$$(n=3, 4, 5, \ldots)$$

и, следовательно,

$$y=\lim_{n \to \infty} y_n=1-x+2[e^{-x}-(1-x)]=2e^{-x}-(1-x),$$

причем сходимость равномерна на любом отрезке $[0, a]$.

Заметим, что дифференциальное уравнение (10) — линейное, с постоянными коэффициентами, и поэтому точное решение, удовлетворяющее заданному начальному условию, может быть найдено элементарно.

Нетрудно также развить метод последовательных приближений для системы дифференциальных уравнений

$$\frac{d\boldsymbol{y}}{dx}=\boldsymbol{f}(x, \boldsymbol{y}), \tag{12}$$

где

$$\boldsymbol{y}(x_0)=\boldsymbol{y}_0. \tag{13}$$

Записывая векторное уравнение (12) в интегральной форме, будем иметь

$$\boldsymbol{y}=\boldsymbol{y}_0+\int_{x_0}^{x} \boldsymbol{f}(x, \boldsymbol{y})\, dx, \tag{14}$$

где под интегралом от вектор-функции

$$\boldsymbol{f}=\begin{bmatrix} f_1 \\ \cdot \\ \cdot \\ \cdot \\ f_n \end{bmatrix}$$

понимается вектор

$$\int_{x_0}^{x} \boldsymbol{f}\, dx = \begin{bmatrix} \int_{x_0}^{x} f_1\, dx \\ \cdot \\ \cdot \\ \cdot \\ \int_{x_0}^{x} f_n\, dx \end{bmatrix}.$$

Последовательные приближения $\boldsymbol{y}^{(p)}$ $(p = 1, 2, \ldots)$ определяются по формуле

$$\boldsymbol{y}^{(p)} = \boldsymbol{y}_0 + \int_{x_0}^{x} \boldsymbol{f}(x, \boldsymbol{y}^{(p-1)})\, dx, \tag{15}$$

причем обычно полагают

$$\boldsymbol{y}^{(0)} \equiv \boldsymbol{y}_0.$$

Этот метод годится также для дифференциального уравнения n-го порядка, если его записать в виде системы.

Пример 2. Построить несколько последовательных приближений для решения системы

$$\frac{dy_1}{dx} = x + y_1 y_2, \qquad \frac{dy_2}{dx} = x^2 - y_1^2,$$

удовлетворяющего начальным условиям $y_1(0) = 1$, $y_2(0) = 0$.

Решение. Имеем

$$y_1 = 1 + \int_0^x (x + y_1 y_2)\, dx, \qquad y_2 = \int_0^x (x^2 - y_1^2)\, dx.$$

Отсюда, полагая $y_1^{(0)} = 1$, $y_2^{(0)} = 0$, получаем

$$y_1^{(1)} = 1 + \int_0^x (x + 0)\, dx = 1 + \frac{x^2}{2},$$

$$y_2^{(1)} = \int_0^x (x^2 - 1)\, dx = -x + \frac{x^3}{3},$$

$$y_1^{(2)} = 1 + \int_0^x \left[x + \left(1 + \frac{x^2}{2}\right)\left(-x + \frac{x^3}{3}\right)\right] dx = 1 - \frac{x^4}{24} + \frac{x^6}{36},$$

$$y_2^{(2)} = \int_0^x \left[x^2 - \left(1 + x^2 + \frac{x^4}{4}\right)\right] dx = -x - \frac{x^5}{20}$$

и т. д.

§ 4. Метод численного интегрирования

Последовательные приближения находятся при помощи квадратур. Может случиться, что эти квадратуры не выражаются в элементарных функциях. В таком случае приходится прибегать к приближенным методам интегрирования.

Приведем один из приемов [14] численного нахождения последовательных приближений для решения уравнения

$$y' = f(x, y) \tag{1}$$

при начальном условии $y(x_0) = y_0$.

Пусть $x_i = x_0 + ih$ $(i = 0, 1, 2, \ldots, m)$ — равноотстоящие значения аргумента x с некоторым шагом h и

$$y_i = y(x_i), \quad y_i' = f(x_i, y_i).$$

Из уравнения (1) имеем

$$y_i - y_0 = \int_{x_0}^{x_i} y'\,dx = h\int_0^i y'\,dq, \tag{2}$$

где $q = \frac{x - x_i}{h}$. Отсюда, применяя одну из квадратурных формул, приближенно получим

$$y_i - y_0 = h\sum_{j=0}^{m} A_j^{(i)} y_j'. \tag{3}$$

Коэффициенты $A_j^{(i)}$ обычно определяются из того условия, чтобы формула (3) являлась точной для всех целых полиномов степени не выше $m+1$. Для этого необходимо и достаточно, чтобы эта формула была справедливой для функций $y = x, x^2, \ldots, x^m, x^{m+1}$ при выборе точек $x_0 = 0$, $x_1 = h$, $x_2 = 2h$, $\ldots$, $x_m = mh$.

Подставляя эти значения в формулу (3), получим для определения $m+1$ неизвестных $A_j^{(i)}$ $(j = 0, 1, 2, \ldots, m)$ при каждом фиксированном i $(i = 1, 2, \ldots)$ систему $m+1$ уравнений

$$i^k = k\sum_{j=0}^{m} A_j^{(i)} j^{(k-1)} \qquad (k = 1, 2, \ldots, m+1), \tag{4}$$

причем

$$\sum_{j=0}^{m} A_j^{(i)} = i.$$

Например, при $m=4$, используя указанный прием, получаем следующие формулы:

$$\left.\begin{aligned} y_1-y_0&=\frac{h}{720}(251y_0'+646y_1'-264y_2'+106y_3'-19y_4'),\\ y_2-y_0&=\frac{h}{90}(29y_0'+124y_1'+24y_2'+4y_3'-y_4'),\\ y_3-y_0&=\frac{h}{80}(27y_0'+102y_1'+72y_2'+42y_3'-3y_4'),\\ y_4-y_0&=\frac{4h}{90}(7y_0'+32y_1'+12y_2'+32y_3'+7y_4'). \end{aligned}\right\} \quad (5)$$

Если $y\in C^{(7)}[x_0, x_4]$, то соответствующие остаточные члены будут равны [14]:

$$R_1=\frac{3}{160}h^6y^{(6)}(\xi_1), \qquad R_3=\frac{3}{160}h^6y^{(6)}(\xi_3),$$

$$R_2=\frac{1}{90}h^6y^{(6)}(\xi_2), \qquad R_4=\frac{-8}{945}h^7y^{(7)}(\xi_4),$$

где $\xi_i\in(x_0, x_4)$ $\quad (i=1, 2, 3, 4)$.

Для вычисления значений последовательных приближений $y^{(p+1)}(x_i)=y_i^{(p+1)}$ $(p=0, 1, 2, \ldots)$ положим

$$y_i^{(p+1)}-y_0=h\sum_{j=0}^{m}A_j^{(i)}\,y_j'^{(p)}, \quad (6)$$

где $y_j'^{(0)}=f(x_j, y^{(0)})$, $y_j'^{(p)}=f(x_j, y_j^{(p)})$ $(p=1, 2, \ldots)$, $A_j^{(i)}$ — коэффициенты, определяемые из системы (4), причем $y^{(0)}$ — известное начальное приближение; в частности, можно принять $y^{(0)}=y_0$.

Пример 1. Численно найти несколько приближений решения уравнения [14]

$$y'=\sin x+\cos y, \quad y(0)=0. \quad (7)$$

Решение. Точное вычисление последовательных приближений $y^{(n)}$ $(n=1, 2, \ldots)$ приводит к неберущимся квадратурам. Поэтому для нахождения приближений применим численный метод.

Выберем шаг $h=0{,}2$ и, пользуясь формулой (6), определим значения последовательных приближений $y^{(n)}$ $(n=1, 2, 3, 4, 5)$, например, в пяти точках $(m=4)$: $x=0$; 0,2; 0,4; 0,6; 0,8.

За начальное приближение примем $y^{(0)}=y(0)+xy'(0)=x$. Отсюда $y_i'^{(0)}=\sin x_i+\cos x_i$. Используя этот результат, по формулам (6), где коэффициенты $A_j^{(i)}$ берутся из формул (5), находим значения $y_i^{(1)}$ первого приближения и, следовательно, получаем

$$y_i'^{(1)}=\sin x_i+\cos y_i^{(1)}.$$

Этот процесс повторяем нужное число раз. Результаты вычислений приведены в таблице 33.

Таблица 33

Численное нахождение последовательных приближений интеграла дифференциального уравнения (7)

x	y	$\sin x$	$\cos y$	y'	Приближение
0	0	0	1	1	$y_0(x)$
0,2	0,2	0,19867	0,98007	1,17874	
0,4	0,4	0,38942	0,92106	1,31048	
0,6	0,6	0,56464	0,82534	1,38998	
0,8	0,8	0,71736	0,69671	1,41407	
0	0	0	1	1	$y_1(x)$
0,2	0,21860	0,19867	0,97620	1,17487	
0,4	0,46836	0,38942	0,80231	1,28173	
0,6	0,73931	0,56464	0,73893	1,30357	
0,8	1,02065	0,71736	0,52281	1,24017	
0	0	0	1	1	$y_2(x)$
0,2	0,21839	0,19867	0,97625	1,17492	
0,4	0,46538	0,38942	0,89365	1,28307	
0,6	0,72538	0,56464	0,74825	1,31289	
0,8	0,98108	0,71736	0,55612	1,27348	
0	0	0	1	1	$y_3(x)$
0,2	0,21840	0,19867	0,97625	1,17492	
0,4	0,46547	0,38942	0,89361	1,28303	
0,6	0,72636	0,56464	0,74760	1,31224	
0,8	0,98596	0,71736	0,55206	1,26942	
0	0	0	1	1	$y_4(x)$
0,2	0,21840	0,19867	0,97625	1,17492	
0,4	0,46547	0,38942	0,89361	1,28303	
0,6	0,72631	0,56464	0,74763	1,31227	
0,8	0,98552	0,71736	0,55243	1,26979	
0	0			1	$y_5(x)$
0,2	0,21840			1,17492	
0,4	0,46547			1,28303	
0,6	0,72631			1,31227	
0,8	0,98555			1,26976	

Можно принять $y(x) \approx y_5(x)$ с точностью до 10^{-5} (за исключением $x=0{,}8$). В случае необходимости таблицу можно продолжить, приняв последнюю строку приближения $y_5(x)$ за первую строку новой таблицы и так далее.

Следует отметить, что лучшую сходимость процесса численных приближений получим, если вместо формул (3) будем пользоваться центральными симметричными формулами, которые определяют значе-

ния последовательных приближений в точках, лежащих симметрично по обе стороны от начальной точки.

Для системы пяти равноотстоящих точек x_{-2}, x_{-1}, x_0, x_1, x_2 эти формулы имеют вид [14]:

$$\left.\begin{aligned}
y_{-2}-y_0 &= -\frac{h}{90}(29y'_{-2}+124y'_{-1}+24y'_0+4y'_1-y'_2)+\frac{41h^6y^{(6)}(\xi)}{3600},\\
y_{-1}-y_0 &= -\frac{h}{720}(-19y'_{-2}+346y'_{-1}+\\
&\qquad +456y'_0-74y'_1+11y'_2)+\frac{11h^6y^{(6)}(\xi)}{1440},\\
y_1-y_0 &= \frac{h}{720}(11y'_{-2}-74y'_{-1}+456y'_0+346y'_1-19y'_2)+\frac{11h^6y^{(6)}(\xi)}{1440},\\
y_2-y_0 &= \frac{h}{90}(-y'_{-2}+4y'_{-1}+24y'_0+124y'_1+29y'_2)+\frac{41h^6y^{(6)}(\xi)}{3600},
\end{aligned}\right\} \quad (8)$$

где $h=x_{i+1}-x_i$ и значения производных $y^{(6)}$ берутся в некоторых промежуточных точках, причем предполагается, что $y\in C^{(6)}[x_{-2}, x_2]$.

Пример 2. Проинтегрировать на отрезке $[-1{,}2;\ -0{,}8]$ уравнение [14]

$$y'=x^2-y^2$$

при начальном условии $y(-1)=0$.

Решение. Используем метод последовательных приближений. Положим $y_0=0$, тогда

$$y_1=\int_{-1}^{x} t^2\,dt=\frac{x^3}{3}+\frac{1}{3},$$

$$y_2=\int_{-1}^{x}\left[t^2-\frac{1}{9}(t^6+2t^3+1)\right]dt=\frac{1}{126}(33-14x+42x^3-7x^4-2x^7),$$

$$y_3=\int_{-1}^{x}\left[t^2-\frac{1}{15\,876}(33-14t+42t^2-7t^4-2t^7)^2\right]dt=$$

$$=\frac{1}{2619540}(724780-179685x+76230x^2+862400x^3-$$
$$-114345x^4+54054x^5-5390x^6-41580x^7-14850x^8-$$
$$-1925x^9+2520x^{11}-385x^{12}-44x^{15}).$$

Отсюда видно, к каким громоздким вычислениям может привести метод последовательных приближений.

Заметим кстати, что четвертое приближение y_4 выразится полиномом 31-й степени, а y_5 — полиномом 63-й степени.

Для упрощения вычислений применим процесс численного нахождения приближений, вполне аналогичный разобранному выше, используя для этого формулы (8). За начальное приближение примем

$$y^{(0)}=x+1.$$

Результаты вычислений приведены в таблице 34.

Таблица 34

Численное нахождение последовательных приближений интеграла дифференциального уравнения с помощью центральных формул

x	y	y'	Приближение	x	y	y'	Приближение
−1,2	−0,2	1,40	1-е приближение	−1,2	−0,23914	1,38281	3-е приближение
−1,1	−0,1	1,20		−1,1	−0,10995	1,19791	
−1,0	0	1,00		−1,0	0	1,00000	
−0,9	0,1	0,80		−0,9	0,09005	0,80189	
−0,8	0,2	0,60		−0,8	0,16074	0,61416	
−1,2	−0,240	1,3824	2-е приближение	−1,2	−0,23916	1,38280	4-е приближение
−1,1	−0,110	1,1979		−1,1	−0,10995	1,19791	
−1,0	0	1,0000		−1,0	0	1,00000	
−0,9	0,090	0,8019		−0,9	0,09005	0,80189	
−0,8	0,160	0,6144		−0,8	0,16073	0,61417	

Мы видим, что 3-е и 4-е приближения совпадают между собой с точностью до 4-го десятичного знака.

Указанный способ численного нахождения последовательных приближений решения дифференциального уравнения с использованием квадратурных формул является весьма эффективным для получения нескольких начальных значений искомого решения, необходимых для начала вычислений в ряде методов приближенного интегрирования (см. ниже).

§ 5. Метод Эйлера

Пусть дано дифференциальное уравнение

$$y' = f(x, y) \tag{1}$$

с начальным условием $y(x_0) = y_0$.

Выбрав достаточно малый шаг h, построим систему равноотстоящих точек

$$x_i = x_0 + ih \qquad (i = 0, 1, 2, \ldots). \tag{2}$$

Искомую интегральную кривую $y = y(x)$, проходящую через точку $M_0(x_0, y_0)$, приближенно заменим (рис. 28) ломаной $M_0M_1M_2 \ldots$ [3], [4] с вершинами $M_i(x_i, y_i)$ $(i = 0, 1, 2, \ldots)$, звенья которой M_iM_{i+1} прямолинейны между прямыми $x = x_i$, $x = x_{i+1}$ и имеют подъем

$$\frac{y_{i+1} - y_i}{h} = f(x_i, y_i) \tag{3}$$

(так называемая *ломаная Эйлера*).

Таким образом, звенья M_iM_{i+1} ломаной Эйлера в каждой вершине M_i имеют направление $y_i = f(x_i, y_i)$, совпадающее с напра-

влением интегральной кривой уравнения (1), проходящей через точку M_i.

Из формулы (3) вытекает, что значения y_i могут быть определены (*метод Эйлера*) по формулам

$$y_{i+1} = y_i + \Delta y_i, \quad \Delta y_i = hf(x_i, y_i) \quad (i = 0, 1, 2, \ldots). \tag{4}$$

Для геометрического построения ломаной Эйлера выберем полюс $P(-1, 0)$ и на оси ординат отложим отрезок $OA_0 = f(x_0, y_0)$ (рис. 28).

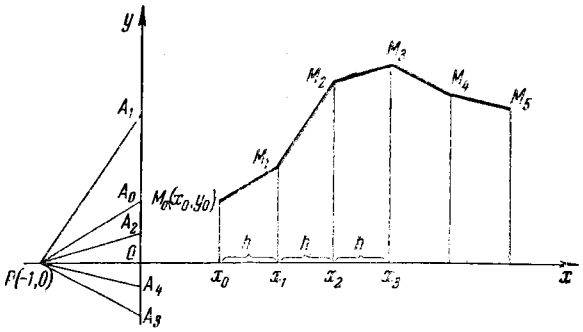

Рис. 28.

Очевидно, угловой коэффициент луча PA_0 будет равен $f(x_0, y_0)$; поэтому, чтобы получить первое звено ломаной Эйлера, достаточно из точки M_0 провести прямую M_0M_1, параллельную лучу PA_0, до пересечения с прямой $x = x_1$ в некоторой точке $M_1(x_1, y_1)$.

Приняв точку $M_1(x_1, y_1)$ за исходную, откладываем на оси ординат отрезок $OA_1 = f(x_1, y_1)$ и через точку M_1 проводим прямую $M_1M_2 \parallel PA_1$ до пересечения в точке M_2 с прямой $x = x_2$ и т. д.

Метод Эйлера является простейшим численным методом интегрирования дифференциального уравнения. Недостатки его:

1) малая точность;

2) систематическое накопление ошибок.

Можно доказать [1], что если правая часть $f(x, y)$ уравнения (1) непрерывна, то последовательность ломаных Эйлера при $h \to 0$ на достаточно малом отрезке $[x_0, x_0 + H]$ равномерно стремится к искомой интегральной кривой $y = y(x)$.

Метод Эйлера легко распространяется на системы дифференциальных уравнений.

Пример 1. Применяя метод Эйлера, составить на отрезке [0, 1] таблицу значений интеграла дифференциального уравнения

$$y' = \frac{xy}{2}, \tag{5}$$

удовлетворяющего начальному условию $y(0) = 1$, выбрав шаг $h = 0{,}1$.

Решение. Результаты вычислений приведены в таблице 35. Для сравнения в последнем столбце помещены значения точного решения $y=e^{x^2/4}$. Из приведенной таблицы видно, что абсолютная погрешность значения y_{10} составляет $\varepsilon_{10}=0{,}0361$. Отсюда относительная погрешность примерно равна 3%.

Для сравнения приводим график точного решения (выделенный жирной линией) и соответствующую ломаную Эйлера $M_0M_1M_2\ldots$ (рис. 29).

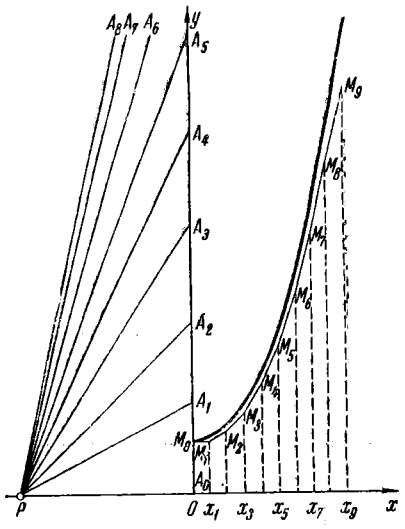

Рис. 29.

Метод Эйлера, вообще говоря, обладает малой точностью и дает сравнительно удовлетворительные результаты (в смысле погрешности) лишь при малых значениях h. Это обстоятельство понятно, так как по существу метод Эйлера заключается в том, что интеграл дифференциального уравнения (1) на каждом частичном отрезке $[x_i,\ x_{i+1}]$ представляется двумя членами ряда Тейлора

$$y(x_i+h)=y(x_i)+hy'(x_i)$$
$$(i=0,\ 1,\ 2,\ \ldots),$$

т. е. для этого отрезка имеется погрешность порядка h^2.

Таблица 35

Интегрирование дифференциального уравнения (5) методом Эйлера

i	x	y	$f(x,\ y)=\frac{xy}{2}$	$\Delta y=0{,}1f(x,\ y)$	Точное значение $y=e^{x^2/4}$
0	0	1	0	0	1
1	0,1	1	0,05	0,005	1,0025
2	0,2	1,005	0,1005	0,0101	1,0100
3	0,3	1,0151	0,1523	0,0152	1,0227
4	0,4	1,0303	0,2067	0,0206	1,0408
5	0,5	1,0509	0,2627	0,0263	1,0645
6	0,6	1,0772	0,3232	0,0323	1,0942
7	0,7	1,1095	0,3883	0,0388	1,1303
8	0,8	1,1483	0,4593	0,0459	1,1735
9	0,9	1,1942	0,5374	0,0537	1,2244
10	1,0	1,2479			1,2840

Кроме того, при вычислении значений на следующем отрезке исходные данные не являются точными и содержат погрешности, зависящие от неточности предшествующих вычислений.

В следующем параграфе будут рассмотрены некоторые приемы уточнения метода Эйлера.

§ 6. Модификации метода Эйлера

Рассмотрим снова дифференциальное уравнение

$$y' = f(x, y) \tag{1}$$

с начальным условием

$$y(x_0) = y_0. \tag{2}$$

Выбрав шаг h, положим

$$x_i = x_0 + ih \quad (i = 0, 1, 2, \ldots).$$

Согласно методу Эйлера последовательные значения искомого решения вычисляются по приближенной формуле

$$y_{i+1} = y_i + hf_i, \quad f_i = f(x_i, y_i). \tag{3}$$

Более точным является усовершенствованный метод ломаных [18], при котором сначала вычисляют промежуточные значения

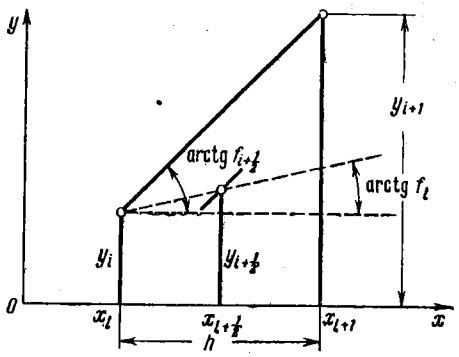

Рис. 30.

$$x_{i+\frac{1}{2}} = x_i + \frac{h}{2}, \quad y_{i+\frac{1}{2}} = y_i + \frac{h}{2} f_i$$

и находят значение направления поля интегральных кривых в средней точке $\left(x_{i+\frac{1}{2}}, y_{i+\frac{1}{2}}\right)$, т. е. $f_{i+\frac{1}{2}} = f\left(x_{i+\frac{1}{2}}, y_{i+\frac{1}{2}}\right)$, а затем полагают (рис. 30)

$$y_{i+1} = y_i + hf_{i+\frac{1}{2}}. \tag{4}$$

Другой модификацией метода Эйлера является усовершенствованный метод Эйлера—Коши [18], при котором сначала определяется «грубое приближение» решения

$$\tilde{y}_{i+1} = y_i + hf_i,$$

исходя из которого находится направление поля интегральных кривых

$$\tilde{f}_{i+1} = f(x_{i+1}, \tilde{y}_{i+1}).$$

Затем приближенно полагают

$$y_{i+1} = y_i + h\frac{f_i + \tilde{f}_{i+1}}{2} \qquad (5)$$

(рис. 31).

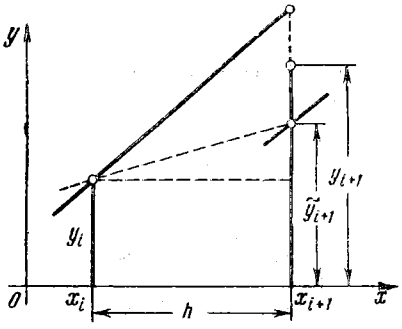

Рис. 31.

Пример 1 Первым и вторым усовершенствованными методами Эйлера проинтегрировать уравнение

$$y' = y - \frac{2x}{y}, \quad y(0) = 1 \qquad (6)$$

на отрезке [0, 1].

Решение. Примем шаг $h = 0{,}2$ и $f(x, y) = y - 2xy^{-1}$.

Приближенные значения искомого решения $y = y(x)$, определенные с помощью усовершенствованного метода ломаных (4), помещены в таблице 36.

Таблица 36

Интегрирование дифференциального уравнения (6) усовершенствованным методом ломаных

i	x_i	y_i	$\frac{h}{2} f_i$	$x_{i+\frac{1}{2}}$	$y_{i+\frac{1}{2}}$	$\Delta y_i = hf_{i+\frac{1}{2}}$
0	0,	1	0,1	0,1	1,1	0,1836
1	0,2	1,1836	0,0846	0,3	1,2682	0,1590
2	0,4	1,3426	0,0747	0,5	1,4173	0,1424
3	0,6	1,4850	0,0677	0,7	1,5527	0,1302
4	0,8	1,6152	0,0625	0,9	1,6777	0,1210
5	1,0	1,7362				

В таблице 37 приведены результаты вычислений интеграла уравнения (6) усовершенствованным методом Эйлера—Коши, причем шаг сохранен прежний: $h = 0{,}2$.

Для сравнения приводим точное решение $y=\sqrt{2x+1}$, откуда $y(1)=\sqrt{3}=1{,}73205\ldots$

Таблица 37

Интегрирование дифференциального уравнения (6) усовершенствованным методом Эйлера — Коши

i	x_i	y_i	$\frac{h}{2}f_i$	x_{i+1}	$\tilde{y}_{i+1}$	$\frac{h}{2}\tilde{f}_{i+1}$	$\Delta y_i=\frac{h}{2}(f_i+\tilde{f}_{i+1})$
0	0	1	0,1	0,2	1,2	0,0867	0,1867
1	0,2	1,1867	0,0850	0,4	1,3566	0,0767	0,1617
2	0,4	1,3484	0,0755	0,6	1,4993	0,0699	0,1454
3	0,6	1,4938	0,0690	0,8	1,618	0,0651	0,1341
4	0,8	1,6279	0,0645	1,0	1,7569	0,0618	0,1263
5	1,0	1,7542					

Усовершенствованный метод Эйлера — Коши можно еще более уточнить, применяя итерационную обработку [19] каждого значения y_i. А именно, исходя из грубого приближения

$$y_{i+1}^{(0)}=y_i+hf(x_i,\ y_i),$$

построим итерационный процесс

$$y_{i+1}^{(k)}=y_i+\frac{h}{2}\left[f(x_i,\ y_i)+f(x_{i+1},\ y_i^{(k-1)})\right]\quad (k=1,\ 2,\ \ldots). \qquad (7)$$

Итерацию продолжаем до тех пор, пока некоторые два последовательных приближения $y_{i+1}^{(m)}$ и $y_{i+1}^{(m+1)}$ не совпадут между собой в соответствующих десятичных знаках. После этого полагаем $y_{i+1}\approx\overline{y}_{i+1}^{(m)}$, где $\overline{y}_{i+1}^{m}$ — общая часть приближений $y_{i+1}^{(m)}$ и $y_{i+1}^{(m+1)}$.

Если алгоритм уточнения численного значения y_i искомого решения после трех-четырех итераций не приводит к совпадению требуемого числа десятичных знаков, то следует уменьшить шаг вычисления h.

Отметим в заключение, что метод Эйлера с итерационной обработкой ординат дает на каждом шаге погрешность порядка h^3 и нередко применяется в вычислительной практике.

Пример 2. Применяя метод итерационной обработки, с точностью до четырех совпадающих десятичных знаков найти значение $y(0{,}1)$ интеграла дифференциального уравнения

$$y'=x+y;\quad y(0)=1.$$

Решение. Примем шаг $h=0{,}05$. Применяя итерационный процесс (7) и учитывая, что $f(x_1, y)=0{,}05+y$, последовательно будем иметь:

$$y_1^{(0)}=1+0{,}05\cdot 1=1{,}05,$$

$$y_1^{(1)}=1+0{,}05\cdot\frac{1+1{,}10}{2}=1{,}0525.$$

$$y_1^{(2)}=1+0{,}05\cdot\frac{1+1{,}1025}{2}=1{,}05256,$$

$$y_1^{(3)}=1+0{,}05\cdot\frac{1+1{,}10256}{2}=1{,}05256.$$

Следовательно, удержав запасной знак, можно положить $y_1=1{,}0526$.

Аналогично, приняв $x_1=0{,}05$ и $y_1=1{,}0526$ за исходные данные и принимая во внимание, что $f(x_2, y)=0{,}1+y$, с помощью того же итерационного процесса (7) получим:

$$y_2^{(0)}=1{,}0526+0{,}05\cdot 1{,}1026=1{,}1077,$$

$$y_2^{(1)}=1{,}0526+0{,}05\cdot\frac{1{,}1026+1{,}2077}{2}=1{,}1104,$$

$$y_2^{(2)}=1{,}0526+0{,}05\cdot\frac{1{,}1026+1{,}2104}{2}=1{,}1104.$$

Отсюда $y_2=1{,}1104$.

Для сравнения приводим точное значение

$$y(0{,}1)=2e^{0{,}1}-1{,}1=1{,}1102.$$

Пример 3. С помощью итерационной обработки уточнить значения y_i ($i=0, 1, \ldots, 10$) интеграла уравнения (§ 5, пример 1)

$$y'=\frac{xy}{2};\quad y(0)=1.$$

Решение. После трех-четырех пересчетов по формулам (7) при $h=0{,}1$ мы приходим к совпадению четвертого знака после запятой в значениях ординат $y_i^{(k)}$ и $y_i^{(k+1)}$. Результаты этих вычислений помещены в таблице 38.

Таблица 38

Интегрирование дифференциального уравнения методом Эйлера с последующей итерацией значений

i	x	y	$f(x, y)=\frac{xy}{2}$	$hf(x, y)$	Точное решение $y=e^{x^2/4}$
0	0	1	0	0	1
1	0,1	1,0025	0,0501	0,0050	1,0025
2	0,2	1,0100	0,1010	0,0101	1,0100
3	0,3	1,0227	0,1534	0,0153	1,0227
4	0,4	1,0408	0,2082	0,0208	1,0408
5	0,5	1,0646	0,2661	0,0266	1,0645
6	0,6	1,0943	0,3283	0,0328	1,0942
7	0,7	1,1305	0,3957	0,0396	1,1303
8	0,8	1,1738	0,4695	0,0470	1,1735
9	0,9	1,2248	0,5512	0,0551	1,2244
10	1,0	1,2812			1,2840

Из таблицы видно, что предельная абсолютная погрешность приближенного решения на отрезке [0,1] меньше $3\cdot10^{-3}$, т. е. предельная относительная погрешность составляет примерно 0,2%.

Метод Эйлера и его модификации являются простейшими представителями *конечно-разностных* методов (*шаговых методов*) для приближенного решения обыкновенных дифференциальных уравнений вида

$$y_k' = f_k(x, y_1, y_2, \ldots, y_n) \qquad (k = 1, 2, \ldots, n) \tag{8}$$

при заданных начальных условиях

$$y_k(x_0) = y_k^{(0)} \qquad (k = 1, 2, \ldots, n). \tag{9}$$

При применении конечно-разностного метода искомое решение $y_k = y_k(x)$ $(k = 1, 2, \ldots, n)$ последовательно строится на системе точек (узлов) $x_i = x_0 + ih$ $(i = 0, 1, 2, \ldots)$, где h — выбранный шаг. Процесс вычислений расчленяется на однообразно повторяющиеся циклы, каждый из которых обеспечивает переход от значения $y_k(x_i)$ к значению $y_k(x_{i+1})$, начиная с начального $y_k^{(0)}$. Поэтому схема вычислений, вообще говоря, легко программируется и удобна для реализации на электронно-счетных машинах. Если правые части f_k системы (8) сложны, то требуются специальные подпрограммы для подсчета $y_k'(x_i)$.

§ 7. Метод Рунге — Кутта

Пусть дано дифференциальное уравнение первого порядка

$$y' = f(x, y) \tag{1}$$

с начальным условием $y(x_0) = y_0$.

Выберем шаг h и для краткости введем обозначения $x_i = x_0 + ih$ и $y_i = y(x_i)$ $(i = 0, 1, 2, \ldots)$.

Рассмотрим числа:

$$\left.\begin{aligned} k_1^{(i)} &= hf(x_i, y_i), & k_3^{(i)} &= hf\left(x_i + \frac{h}{2}, y_i + \frac{k_2^{(i)}}{2}\right), \\ k_2^{(i)} &= hf\left(x_i + \frac{h}{2}, y_i + \frac{k_1^{(i)}}{2}\right), & k_4^{(i)} &= hf(x_i + h, y_i + k_3^{(i)}). \end{aligned}\right\} \tag{2}$$

Согласно обычному методу Рунге—Кутта [18]—[20] последовательные значения y_i искомой функции y определяются по формуле

$$y_{i+1} = y_i + \Delta y_i,$$

где

$$\Delta y_i = \frac{1}{6}(k_1^{(i)} + 2k_2^{(i)} + 2k_3^{(i)} + k_4^{(i)}) \qquad (i = 0, 1, 2, \ldots). \tag{3}$$

Можно доказать, что погрешность этого метода на каждом шаге есть величина порядка h^5 [18], в предположении, что $f(x, y) \in C^{(5)}$. Получены также формулы типа Рунге—Кутта с иными порядками точности [18], [24].

Для вычисления по формуле (3) удобно пользоваться схемой, приведенной в таблице 39.

Таблица 39

Схема метода Рунге — Кутта

i	x	y	$k=hf(x, y)$	Δy
0	x_0	y_0	$k_1^{(0)}$	$k_1^{(0)}$
	$x_0+\frac{h}{2}$	$y_0+\frac{k_1^{(0)}}{2}$	$k_2^{(0)}$	$2k_2^{(0)}$
	$x_0+\frac{h}{2}$	$y_0+\frac{k_2^{(0)}}{2}$	$k_3^{(0)}$	$2k_3^{(0)}$
	x_0+h	$y_0+k_3^{(0)}$	$k_4^{(0)}$	$k_4^{(0)}$
—	—	—	—	$\frac{1}{6}\Sigma=\Delta y_0$
1	x_1	y_1	...	...

Эффективная оценка погрешности метода Рунге—Кутта затруднительна (см. [18]). Поэтому для определения правильности выбора шага h на практике обычно на каждом этапе из двух шагов применяют двойной пересчет [26]. А именно, исходя из текущего верного значения $y(x_i)$, вычисляют величину $y(x_i+2h)$ двумя способами: один раз с шагом h, а другой раз с двойным шагом $H=2h$. Если расхождение полученных значений не превышает допустимой погрешности, то шаг h для данного этапа выбран правильно и полученное с его помощью значение можно принять за $y(x_i+2h)$. В противном случае шаг уменьшают в два раза. Такого рода вычислительную схему легко запрограммировать для работы на электронно-счетных машинах [26]. Употребляется также приближенная оценка погрешности (см. § 13).

Метод Рунге—Кутта обладает значительной точностью и, несмотря на свою трудоемкость, широко используется при численном решении дифференциальных уравнений с помощью электронных вычислительных машин. Кроме того, важным преимуществом этого метода является возможность применения «переменного шага».

Заметим, что для начала вычислений по методу Рунге—Кутта не нужно строить начальный отрезок (см. § 8).

Пример 1. Методом Рунге—Кутта вычислить на отрезке [0, 0,5] интеграл дифференциального уравнения

$$y' = x + y, \quad y(0) = 1, \tag{4}$$

приняв шаг $h = 0{,}1$.

Решение. Покажем начало процесса.

Вычисление y_1. Последовательно имеем:

$$k_1^{(0)} = (0+1)\cdot 0{,}1 = 0{,}1,$$

$$k_2^{(0)} = 0{,}05 + (1+0{,}05)\cdot 0{,}1 = 0{,}11,$$

$$k_3^{(0)} = 0{,}05 + (1+0{,}055)\cdot 0{,}1 = 0{,}1105,$$

$$k_4^{(0)} = 0{,}1 + (1+0{,}1105)\cdot 0{,}1 = 0{,}12105.$$

Отсюда

$$\Delta y_0 = \frac{1}{6}(0{,}1 + 2\cdot 0{,}11 + 2\cdot 0{,}1105 + 0{,}12105) = 0{,}1103,$$

и, следовательно,

$$y_1 = y_0 + \Delta y_0 = 1 + 0{,}1103 = 1{,}1103.$$

Аналогично вычисляются дальнейшие приближения. Результаты вычислений приведены в таблице 40. Таким образом, $y(0{,}5) = 1{,}7974$. Для сравнения приводим точное решение $y = 2e^x - x - 1$, откуда $y(0{,}5) = 2\sqrt{e} - 1{,}5 = 1{,}79744\ldots$

Метод Рунге—Кутта применим также для приближенного решения систем обыкновенных дифференциальных уравнений.

Пусть дана система дифференциальных уравнений

$$\frac{d\boldsymbol{y}}{dx} = \boldsymbol{f}(x, \boldsymbol{y}) \tag{5}$$

и начальные условия $\boldsymbol{y}(x_0) = \boldsymbol{y}_0$.

Задавшись некоторым шагом h и введя стандартные обозначения $x_i = x_0 + ih$ и $\boldsymbol{y}_i = \boldsymbol{y}_i(x_i)$, $\Delta\boldsymbol{y}_i = \boldsymbol{y}_{i+1} - \boldsymbol{y}_i$ при $i = 0, 1, 2, \ldots$, положим:

$$\left.\begin{aligned}
\boldsymbol{k}_1^{(0)} &= h\boldsymbol{f}(x_0, \boldsymbol{y}_0),\\
\boldsymbol{k}_2^{(0)} &= h\boldsymbol{f}\left(x_0 + \frac{h}{2}, \boldsymbol{y}_0 + \frac{\boldsymbol{k}_1^{(0)}}{2}\right),\\
\boldsymbol{k}_3^{(0)} &= h\boldsymbol{f}\left(x_0 + \frac{h}{2}, \boldsymbol{y}_0 + \frac{\boldsymbol{k}_2^{(0)}}{2}\right),\\
\boldsymbol{k}_4^{(0)} &= h\boldsymbol{f}(x_0 + h, \boldsymbol{y}_0 + \boldsymbol{k}_3^{(0)}).
\end{aligned}\right\} \tag{6}$$

Согласно методу Рунге—Кутта $\Delta\boldsymbol{y}_0$ приближенно определяют по формуле

$$\Delta\boldsymbol{y}_0 = \frac{1}{6}\left(\boldsymbol{k}_1^{(0)} + 2\boldsymbol{k}_2^{(0)} + 2\boldsymbol{k}_3^{(0)} + \boldsymbol{k}_4^{(0)}\right), \tag{7}$$

отсюда $\boldsymbol{y}_1 = \boldsymbol{y}_0 + \Delta\boldsymbol{y}_0$.

Таблица 40

Интегрирование дифференциального уравнения (4) методом Рунге — Кутта

i	x	y	$k=0,1\ (x+y)$	Δy
0	0 0,05 0,05 0,1	1 1,05 1,055 1,1105	0,1 0,11 0,1105 0,1210	0,1000 0,2200 0,2210 0,1210
				$\frac{1}{6}\cdot 0,6620=0,1103$
1	0,1 0,15 0,15 0,2	1,1103 1,1708 1,1763 1,2429	0,1210 0,1321 0,1326 0,1443	0,1210 0,2642 0,2652 0,1443
				$\frac{1}{6}\cdot 0,7947=0,1324$
2	0,2 0,25 0,25 0,3	1,2427 1,3149 1,3209 1,3998	0,1443 0,1565 0,1571 0,1700	0,1443 0,3130 0,3142 0,1700
				$\frac{1}{6}\cdot 0,9415=0,1569$
3	0,3 0,35 0,35 0,4	1,3996 1,4846 1,4904 1,5836	0,1700 0,1835 0,1840 0,1984	0,1700 0,3670 0,3680 0,1984
				$\frac{1}{6}\cdot 1,1034=0,1840$
4	0,4 0,45 0,45 0,5	1,5836 1,6828 1,6902 1,7976	0,1984 0,2133 0,2140 0,2298	0,1984 0,4266 0,4280 0,2298
				$\frac{1}{6}\cdot 1,2828=0,2138$
5	0,5	1,7974		

Далее, приняв (x_1, y_1) за исходные данные и повторяя тот же процесс, находим y_2. Аналогично вычисляются y_i $(i=3, 4, \ldots)$.

Пример 2. Методом Рунге—Кутта проинтегрировать уравнение колебаний маятника в сопротивляющейся среде (см. § 1, пример 1)

$$\frac{d^2\Theta}{dt^2}+0{,}2\frac{d\Theta}{dt}+10\sin\Theta=0 \tag{8}$$

при начальных условиях $\Theta(0)=0{,}3$, $\dot{\Theta}(0)=0$ $\left(\dot{\Theta}=\frac{d\Theta}{dt}\right)$.

Таблица 41

Интегрирование системы дифференциальных уравнений (9) методом Рунге — Кутта

i	t	Θ	ω	$k^{(1)}=0{,}1\dot{\omega}$	$k^{(2)}=0{,}1\,\dot{\omega}$	$\Delta\Theta$	$\Delta\omega$
0	0	0,3	0	0	—0,2955	0,	—0,2955
	0,05	0,3000	—0,1478	—0,0148	—0,2926	—0,0296	—0,5852
	0,05	0,2926	—0,1463	—0,0146	—0,2855	—0,0292	—0,5710
	0,1	0,2854	—0,2855	—0,0286	—0,2810	—0,0286	—0,2810
						$-0{,}0874\cdot\frac{1}{6}=$ $=-0{,}0146$	$-1{,}327\cdot\frac{1}{6}=$ $=-0{,}2888$
1	0,1	0,2854	—0,2888	—0,0289	—0,2759	—0,0289	—0,2759
	0,15	0,2710	—0,4267	—0,0427	—0,2592	—0,0854	—0,5184
	0,15	0,2641	—0,4184	—0,0418	—0,2527	—0,0836	—0,5054
	0,2	0,2436	—0,5415	—0,0541	—0,2304	—0,0541	—0,2304
						$-0{,}2520\cdot\frac{1}{6}=$ $=-0{,}0420$	$-1{,}5301\cdot\frac{1}{6}=$ $=-0{,}2550$
2	0,2	0,2434	—0,5438	—0,0544	—0,2301	—0,0544	—0,2301
	0,25	0,2162	—0,6589	—0,0659	—0,2013	—0,1318	—0,4026
	0,25	0,2105	—0,4645	—0,0644	—0,1960	—0,1288	—0,3920
	0,3	0,1790	—0,7398	—0,0740	—0,1633	—0,0740	—0,1633
						$-0{,}3889\cdot\frac{1}{6}=$ $=-0{,}0648$	$-1{,}1895\cdot\frac{1}{6}=$ $=-0{,}1980$
3	0,3	0,1786	—0,7418	—0,0742	—0,1647		

Решение. Полагая $\frac{d\Theta}{dt}=\omega$, запишем уравнение (8) в виде системы дифференциальных уравнений

$$\dot{\Theta}=\omega, \quad \dot{\omega}=-0{,}2\omega-10\sin\Theta, \tag{9}$$

причем $\Theta_0=0{,}3$, $\omega_0=0$. Выберем шаг

$$h=\Delta t=0{,}1$$

и положим

$$\boldsymbol{f}=\begin{bmatrix}\dot{\Theta}\\ \dot{\omega}\end{bmatrix}, \qquad \boldsymbol{k}=\begin{bmatrix}k^{(1)}\\ k^{(2)}\end{bmatrix},$$

где $\dot{\Theta}=\omega$, $\dot{\omega}=-0{,}2\omega-10\sin\Theta$ и компоненты $k^{(1)}$, $k^{(2)}$ определяются из формул (6).

Результаты вычислений по формулам (6) и (7) помещены в таблице 41. Ход дальнейших вычислений понятен из приведенного образца.

§ 8. Метод Адамса

Этот метод численного интегрирования разработан Адамсом в 1855 г. по просьбе известного английского артиллериста Башфорта, занимавшегося внешней баллистикой. Впоследствии этот метод был забыт и вновь открыт в начале века норвежским математиком Штермером. Популяризация метода Адамса и дальнейшее его усовершенствование связаны с именем А. Н. Крылова [10].

Изложим метод Адамса [10], [19], [20] применительно к уравнению первого порядка

$$y'=f(x, y) \tag{1}$$

с начальным условием

$$y(x_0)=y_0. \tag{2}$$

Пусть $x_i\,(i=0, 1, 2, \ldots)$ — система равноотстоящих значений с шагом h и $y_i=y(x_i)$. Очевидно, имеем

$$\Delta y_i=\int_{x_i}^{x_{i+1}} y'\,dx. \tag{3}$$

В силу второй интерполяционной формулы Ньютона (см., например, [13]) с точностью до разностей четвертого порядка получаем

$$y'=y'_i+q\Delta y'_{i-1}+\frac{q(q+1)}{2!}\Delta^2 y'_{i-2}+\frac{q(q+1)(q+2)}{3!}\Delta^3 y'_{i-3}, \tag{4}$$

где $q=\frac{x-x_i}{h}$, или

$$y'=y'_i+q\Delta y'_{i-1}+\frac{q^2+q}{2}\Delta^2 y'_{i-2}+\frac{q^3+3q^2+2q}{6}\Delta^3 y'_{i-3}. \tag{4'}$$

Подставляя выражение (4′) в формулу (3) и учитывая, что $dx = h\,dq$, будем иметь

$$\Delta y_i = h\int_0^1 \left(y'_i + q\Delta y'_{i-1} + \frac{q^2+q}{2}\Delta^2 y'_{i-2} + \frac{q^3+3q^2+2q}{6}\Delta^3 y'_{i-3}\right)dq.$$

Отсюда получаем *экстраполяционную формулу Адамса*

$$\Delta y_i = hy'_i + \frac{1}{2}\Delta(hy'_{i-1}) + \frac{5}{12}\Delta^2(hy'_{i-2}) + \frac{3}{8}\Delta^3(hy'_{i-3}). \qquad (5)$$

Для начала процесса нужны четыре начальных значения y_0, y_1, y_2, y_3, так называемый *начальный отрезок*, который определяют, исходя из начального условия (2), каким-нибудь численным методом. Можно, например, использовать метод Рунге—Кутта или разложение в ряд Тейлора

$$y_i = y(x_0 + ih) = y_0 + y'_0(ih) + \frac{y''_0}{2}(ih)^2 + \ldots,$$

где $i = 1, 2, 3$ (или $i = -1, 1, 2$ с соответствующим изменением нумерации). Зная эти значения, из уравнения (1) можно найти значения производных y'_0, y'_1, y'_2, y'_3 и составить таблицу разностей:

$$\Delta(hy'_0),\ \Delta(hy'_1),\ \Delta(hy'_2),\ \Delta^2(hy'_0),\ \Delta^2(hy'_1),\ \Delta^3(hy'_0). \qquad (6)$$

Дальнейшие значения y_i ($i = 4, 5, \ldots$) искомого решения можно шаг за шагом вычислять по формуле Адамса, пополняя по мере необходимости таблицу разностей (6).

Для контроля рекомендуется [8], [9], вычислив первое приближение для Δy_i по формуле

$$\Delta y_i^{\mathrm{I}} = hy'_i + \frac{1}{2}(h\Delta y'_{i-1}) + \frac{5}{12}\Delta^2(hy'_{i-2}) + \frac{3}{8}\Delta^3(hy'_{i-3}),$$

определить $y_{i+1} = y_i + \Delta y_i^{\mathrm{I}}$, подсчитать конечные разности

$$\Delta(hy_i),\ \Delta^2(hy_{i-1}),\ \Delta^3(hy_{i-2}) \qquad (7)$$

и затем найти второе приближение по более точной формуле (см. § 9)

$$\Delta y_i^{\mathrm{II}} = hy'_i + \frac{1}{2}\Delta(hy'_i) - \frac{1}{12}\Delta^2(hy'_{i-1}) - \frac{1}{24}\Delta^3(hy'_{i-2}). \qquad (8)$$

Если Δy_i^{I} и Δy_i^{II} отличаются лишь на несколько единиц последнего сохраняемого десятичного разряда, то можно положить $\Delta y_i = \Delta y_i^{\mathrm{II}}$, а затем, найдя $\Delta y_{i+1} = y_i + \Delta y_i$, перевычислить конечные разности (7). После этого, строго говоря, следует снова найти Δy_i^{II} по формуле (8). Поэтому шаг h должен быть таким, чтобы этот пересчет был излишним.

На практике шаг h выбирают столь малым, чтобы можно было пренебречь членом $\frac{1}{24}\Delta^3(hy'_{i-2})$ в формуле (8).

Если же расхождение величин Δy_i^{I} и Δy_i^{II} значительно, то следует уменьшить шаг h.

Обычно шаг h уменьшают ровно в два раза. Покажем, как в этом случае, имея до некоторого значения i таблицу величин x_j, y_j, $Y_j = hy'_j$ $(j \leqslant i)$ с шагом $\Delta x_j = h$, можно просто построить таблицу величин $x_i + k\frac{h}{2}$, $y\left(x_i + k\frac{h}{2}\right)$, $\eta_k = \frac{h}{2}y'\left(x_i + k\frac{h}{2}\right)$ $(k = 0, 1, 2, \ldots)$ с шагом $\frac{h}{2}$ [28]. Для краткости введем сокращенные обозначения:

$$x_{i+\frac{k}{2}} = x_i + k\frac{h}{2}, \quad y_{i+\frac{k}{2}} = y\left(x_i + k\frac{h}{2}\right), \quad y'_{i+\frac{k}{2}} = y'\left(x_i + k\frac{h}{2}\right)$$
$$(i = 0, 1, 2, \ldots; \ k = 0, \pm 1, \pm 2, \ldots).$$

На основании формулы (4) будем иметь

$$Y = Y_j + q\Delta Y_{j-1} + \frac{q(q+1)}{2}\Delta^2 Y_{j-2} + \frac{q(q+1)(q+2)}{6}\Delta^3 Y_{j-3}, \tag{9}$$

где $Y = hy'$. Отсюда, полагая $j = i - 2$ и $q = 1/2$ и учитывая, что $x = x_{i-2} + \frac{h}{2} = x_{i-\frac{3}{2}}$, находим

$$Y_{i-\frac{3}{2}} = Y_{i-2} + \frac{1}{2}\Delta Y_{i-3} + \frac{3}{8}\Delta^2 Y_{i-4} + \frac{5}{16}\Delta^3 Y_{i-5}. \tag{10}$$

Аналогично при $j = i - 1$, $q = 1/2$ из формулы (9) получаем, что аргументу $x = x_{i-1} + h/2 = x_{i-\frac{1}{2}}$ соответствует значение

$$Y_{i-\frac{1}{2}} = Y_{i-1} + \frac{1}{2}\Delta Y_{i-2} + \frac{3}{8}\Delta^2 Y_{i-3} + \frac{5}{16}\Delta^3 Y_{i-4}. \tag{11}$$

Что касается значений Y_{i-1} и Y_i, то они имеются в старой таблице. После этого составляем начальный отрезок для новой таблицы:

$$\eta_{-3} = \frac{1}{2}Y_{i-\frac{3}{2}}, \quad \eta_{-2} = \frac{1}{2}Y_{i-1}, \quad \eta_{-1} = \frac{1}{2}Y_{i-\frac{1}{2}}, \quad \eta_0 = \frac{1}{2}Y_i$$

и находим конечные разности:

$$\begin{aligned} \delta\eta_k &= \eta_{k+1} - \eta_k && (k = -3, -2, -1), \\ \delta^2\eta_k &= \delta\eta_{k+1} - \delta\eta_k && (k = -3, -2), \\ \delta^3\eta_k &= \delta^2\eta_{k+1} - \delta^2\eta_k && (k = -3). \end{aligned}$$

Дальше таблица продолжается обычным путем, посредством соответствующей модификации формулы (5):

$$\delta y_{i+\frac{j}{2}} = \eta_j + \frac{1}{2}\delta\eta_{j-1} + \frac{5}{12}\delta^2\eta_{j-2} + \frac{3}{8}\delta^3\eta_{j-3},$$

$$y_{i+\frac{j+1}{2}} = y_{i+\frac{j}{2}} + \delta y_{i+\frac{j}{2}} \qquad (j=0,\ 1,\ 2,\ \ldots).$$

Для работы на электронных счетных машинах формулу Адамса (5) выгодно применять в раскрытом виде. Учитывая, что

$$\Delta y'_{i-1} = y'_i - y'_{i-1},$$
$$\Delta^2 y'_{i-2} = y'_i - 2y'_{i-1} + y'_{i-2},$$
$$\Delta^3 y'_{i-3} = y'_i - 3y'_{i-1} + 3y'_{i-2} - y'_{i-3},$$

после приведения подобных членов имеем

$$y_{i+1} = y_i + \frac{h}{24}(55y'_i - 59y'_{i-1} + 37y'_{i-2} - 9y'_{i-3}),$$

причем $x_{i+1} = x_i + h$.

Пример 1. Методом Адамса найти на отрезке [0, 1] интеграл уравнения

$$y = x + y, \quad y(0) = 1.$$

Решение. Примем шаг $h = 0{,}1$. Для начала процесса используем значения, найденные методом Рунге—Кутта (см. § 7, пример 1), т. е. $y_0 = 1$; $y_1 = 1{,}1103$; $y_2 = 1{,}2427$; $y_3 = 1{,}3996$.

Дальнейшие вычисления располагаем в двух бланках: основном (таблица 42) и вспомогательном (таблица 43).

Таблица 42

Основной бланк для интегрирования дифференциального уравнения методом Адамса

i	x	y	Δy	hy'	$\Delta(hy')$	$\Delta^2(hy')$	$\Delta^3(hy')$	y^*
0	0,	1,		0,1000	210	23	1	1
1	0,1	1,1103		0,1210	333	24	2	1,1103
2	0,2	1,2427		0,1443	257	26	5	1,2428
3	0,3	1,3996	0,1838	0,1700	283	31	2	1,3997
4	0,4	1,5834	0,2137	0,1983	314	33	3	1,5836
5	0,5	1,7971	0,2469	0,2297	347	36	5	1,7974
6	0,6	2,0440	0,2833	0,2644	383	41	3	2,0442
7	0,7	2,3273	0,3235	0,3027	424	44		2,3275
8	0,8	2,6508	0,3682	0,3451	468			2,6511
9	0,9	3,0190	0,4172	0,3919				3,0192
10	1,0	3,4362						3,4366

О. Сергей

Если правая часть дифференциального уравнения сложна, то в основном бланке приходится вводить промежуточные графы [10]. При заполнении вспомогательного бланка используются согласно формуле (5) наклонные строки.

Таблица 43

Вспомогательный бланк для интегрирования дифференциального уравнения методом Адамса

i	3	4	5	6	7	8	9
hy_i	0,1700	0,1983	0,2297	0,2644	0,3027	0,3451	0,3919
$\frac{1}{2}\Delta(hy'_{-1})$	128	142	157	174	192	212	234
$\frac{5}{12}\Delta^2(hy'_{-2})$	10	11	13	14	15	17	18
$\frac{3}{8}\Delta^3(hy'_{-3})$	0	1	2	1	1	2	1
$\Delta y_i = \sum$	0,1838	0,2137	0,2469	0,2833	0,3235	0,3682	0,4172

В последнем столбце таблицы 42 приведены для сравнения точные значения решения $y^* = 2e^x - x - 1$. Отсюда видно, что максимальная ошибка приближенного решения y не превосходит четырех единиц последнего десятичного разряда. Можно было бы уменьшить эту ошибку, применив двойной пересчет по контрольной формуле и введя соответствующие поправки [18].

Метод Адамса [10], [19], [20] легко распространяется на систему дифференциальных уравнений

$$\frac{d\boldsymbol{y}}{dx} = \boldsymbol{f}(x, \boldsymbol{y})$$

при начальных условиях $\boldsymbol{y}(x_0) = \boldsymbol{y}_0$. А именно, имея векторный начальный отрезок $\boldsymbol{y}_0, \boldsymbol{y}_1, \boldsymbol{y}_2, \boldsymbol{y}_3$, дальнейшие значения координат искомой вектор-функции $\boldsymbol{y} = \boldsymbol{y}(x)$ определяем, используя формулу

$$\Delta \boldsymbol{y}_i = h\boldsymbol{y}_i + \frac{1}{2}\Delta(h\boldsymbol{y}_{i-1}) + \frac{5}{12}\Delta^2(h\boldsymbol{y}_{i-2}) + \frac{3}{8}\Delta^3(h\boldsymbol{y}_{i-3}) \quad (i = 4, 5, \ldots).$$

Для численного нахождения решения можно использовать бланки, аналогичные приведенным выше.

Пример 2. Уравнение колебаний маятника имеет вид (см. § 7, пример 2)

$$\frac{d^2\Theta}{dt^2} + 0{,}2\frac{d\Theta}{dt} + 10\sin\Theta = 0, \tag{12}$$

причем $\Theta(0)=0{,}3$, $\dot{\Theta}(0)=0$. Методом Адамса определить угловую скорость маятника при первом прохождении положения равновесия $\Theta=0$.

Решение. Записываем уравнение (12) в виде системы

$$\dot{\Theta}=\omega,\quad \dot{\omega}=-0{,}2\omega-10\sin\Theta.$$

Начальный отрезок для шага $h=\Delta t=0{,}1$, подсчитанный методом Рунге—Кутта, заимствуем из § 7 (пример 2), причем ограничиваемся тремя десятичными знаками:

$$\begin{array}{lll} t=0; & \Theta_0=0{,}300; & \omega_0=0; \\ t=0{,}1; & \Theta_1=0{,}285; & \omega_1=-0{,}389; \\ t=0{,}2; & \Theta_2=0{,}243; & \omega_2=-0{,}544; \\ t=0{,}3; & \Theta_3=0{,}179; & \omega_3=-0{,}742. \end{array}$$

Пользуясь этими данными, заполняем основной бланк (таблица 44) и вспомогательный бланк (таблица 45).

Таблица 44

Основной бланк для решения системы дифференциальных уравнений методом Адамса

i	t	Θ	$\Delta\Theta$	$h\dot{\Theta}$	$\Delta(h\dot{\Theta})$	$\Delta^2(h\dot{\Theta})$
0	0	0,300		0	−29	4
1	0,1	0,285		−0,029	−25	5
2	0,2	0,243		−0,054	−20	8
3	0,3	0,181	−0,082	−0,074	−12	8
4	0,4	0,099	−0,088	−0,086	− 4	
5	0,5	0,11	−0,089	−0,090		
6	0,6	−0,078				

$\Delta^3(h\dot{\Theta})$	ω	$\Delta\omega$	$h\dot{\omega}$	$\Delta(h\dot{\omega})$	$\Delta^2(h\dot{\omega})$	$\Delta^3(h\dot{\omega})$
1	0		−0,296	20	26	− 5
3	−0,289		−0,276	46	21	− 5
0	−0,544		0,230	67	16	−10
	−0,742	−0,122	−0,163	83	6	
	−0,864	−0,033	−0,080	89		
	−0,897	+0,051	+0,009			
	−0,846					

Таким образом, интересующий нас момент времени, для которого отклонение $\Theta(t)=0$, удовлетворяет неравенству $0{,}5<t<0{,}6$.

Для уточнения составим таблицу разностей (таблица 46).

Таблица 45

Вспомогательный бланк для решения системы дифференциальных уравнений методом Адамса

i	3	4	5
$h\dot{\Theta}_i$	—0,074	—0,086	—0,090
$\frac{1}{2}\Delta(h\dot{\Theta}_{i-1})$	— 10	— 6	— 2
$\frac{5}{12}\Delta^2(h\dot{\Theta}_{i-2})$	+ 2	+ 3	+ 3
$\frac{3}{8}\Delta^3(h\dot{\Theta}_{i-3})$	0	+ 1	0
$\Delta\Theta_i = \sum_1$	—0,082	—0,088	—0,089

i	3	4	5
$h\dot{\omega}_i$	—0,163	—0,080	+0,009
$\frac{1}{2}\Delta(h\dot{\omega}_{i-1})$	+ 34	+ 42	+ 44
$\frac{5}{12}\Delta^2(h\dot{\omega}_{i-2})$	+ 9	+ 7	+ 2
$\frac{3}{8}\Delta^3(h\dot{\omega}_{i-3})$	— 12	— 2	— 4
$\Delta\omega_i = \sum_2$	—0,122	—0,033	+0,051

Таблица 46

Конечные разности величин Θ и ω

t	Θ	$\Delta\Theta$	$\Delta^2\Theta$	ω	$\Delta\omega$	$\Delta^2\omega$
0,3	0,181	—82	—6	—0,742	—122	89
0,4	0,099	—88		—0,864	—33	
0,5	0,011			—0,897		

Пусть $q=(t-0{,}5)/0{,}1=10\,(t-0{,}5)$. Применяя вторую интерполяционную формулу Ньютона, будем иметь

$$0=\Theta=0{,}011+q\,(-0{,}088)+\frac{q\,(q+1)}{2}(-0{,}006).$$

Отсюда

$$q=\frac{0{,}011}{0{,}088}-\frac{0{,}003}{0{,}088}\,q\,(q+1)=0{,}125-0{,}034q\,(q+1).$$

Применяя метод последовательных приближений, получим

$$q^{(0)}=0{,}125,$$
$$q^{(1)}=0{,}125-0{,}034\cdot 0{,}102\cdot 1{,}102=0{,}125-0{,}004=0{,}121,$$
$$q^{(2)}=0{,}125-0{,}034\cdot 0{,}098\cdot 1{,}098=0{,}125-0{,}004=0{,}121.$$

Следовательно, можно принять $q=0{,}121$ и $t=0{,}5+0{,}1q=0{,}512$.

Снова пользуясь второй интерполяционной формулой Ньютона, находим

$$\omega=-0{,}897+q\,(-0{,}033)+\frac{q\,(q+1)}{2}\cdot 0{,}089=-0{,}897-0{,}004+0{,}006=-0{,}895.$$

Таким образом, окончательно имеем $\omega=\Theta=-0{,}895$.

§ 9. Метод А. Н. Крылова последовательных сближений

Как видно из предыдущих параграфов, многие методы численного интегрирования дифференциальных уравнений, включая метод Адамса, распадаются на два этапа:

1) нахождение начального отрезка искомого решения, т. е. начало решения или вход в таблицу;

2) вычисление дальнейших значений на основе найденных величин, т. е. продолжение таблицы.

Здесь мы изложим принадлежащий А. Н. Крылову [10], [8], [22] способ построения начального отрезка *методом последовательных сближений*. Этот метод особенно удобен, когда правая часть дифференциального уравнения задана таблично, что, например, имеет место в задачах внешней баллистики.

Для простоты записи ограничимся рассмотрением дифференциального уравнения первого порядка

$$y'=f(x,\ y) \tag{1}$$

с начальным условием $y(x_0)=y_0$.

Выведем сначала ряд вспомогательных формул, полагая

$$y_i=y(x_0+ih) \text{ и } y_i=f(x_i,\ y_i) \quad (i=0,\ 1,\ 2,\ \ldots).$$

В силу формулы Адамса (см. § 8, (5)) имеем

$$\Delta y_i=\eta_i+\frac{1}{2}\Delta\eta_{i-1}+\frac{5}{12}\Delta^2\eta_{i-2}+\frac{3}{8}\Delta^3\eta_{i-3}, \tag{2}$$

где ради краткости введено обозначение $\eta_i=hy_i'=hf(x_i,\ y_i)$.

Формула (2) обычно называется *формулой наклонной строки*, так как в ней используются разности, стоящие на диагонали таблицы разностей. Учитывая, что

$$\Delta\eta_{i-1}=\Delta\eta_i-\Delta^2\eta_{i-1},\quad \Delta^2\eta_{i-2}=\Delta^2\eta_{i-1}-\Delta^3\eta_{i-2},$$

и полагая постоянными конечные разности третьего порядка

$$\Delta^3\eta_{i-3}\approx\Delta^3\eta_{i-2},$$

из формулы (2) будем иметь

$$\Delta y_i=\eta_i+\frac{1}{2}(\Delta\eta_i-\Delta^2\eta_{i-1})+\frac{5}{12}(\Delta^2\eta_{i-1}-\Delta^3\eta_{i-2})+\frac{3}{8}\Delta^3\eta_{i-2}.$$

Отсюда получаем первую вспомогательную формулу — так называемую *первую формулу ломаной строки*

$$\Delta y_i=\eta_i+\frac{1}{2}\Delta\eta_i-\frac{1}{12}\Delta^2\eta_{i-1}-\frac{1}{24}\Delta^3\eta_{i-2}. \tag{3}$$

Далее, учитывая, что $\Delta^2\eta_{i-1}=\Delta^2\eta_i-\Delta^3\eta_{i-1}$ и $\Delta^3\eta_{i-2}\approx\Delta^3\eta_{i-1}$, из формулы (3) выводим вторую вспомогательную формулу — так называемую *вторую формулу ломаной строки*

$$\Delta y_i=\eta_i+\frac{1}{2}\Delta\eta_i-\frac{1}{12}\Delta^2\eta_i+\frac{1}{24}\Delta^3\eta_{i-1}. \tag{4}$$

Наконец, полагая $\Delta^3\eta_{i-1}\approx\Delta^3\eta_i$, получаем *формулу горизонтальной строки*

$$\Delta y_i=\eta_i+\frac{1}{2}\Delta\eta_i-\frac{1}{12}\Delta^2\eta_i+\frac{1}{24}\Delta^3\eta_i. \tag{5}$$

Заметим, что формулу (5) можно получить непосредственно с помощью интегрирования, в пределах от x_i до x_{i+1}, разложения y' по первой интерполяционной формуле Ньютона:

$$y'=y'_i+q\Delta y'_i+\frac{q^2-q}{2}\Delta^2y'_i+\frac{q^3-3q^2+2q}{6}\Delta^3y'_i.$$

Переходим к описанию метода Крылова последовательных сближений.

Первое сближение состоит в вычислении приближенного значения Δy_0 (для ясности это значение и аналогичные ему подчеркнуты одной линией) по одночленной формуле

$$\underline{\Delta y_0}=\eta_0. \tag{6}$$

После этого находим $\underline{y_1}=y_0+\underline{\Delta y_0}$ и составляем разность $\underline{\Delta\eta_0}=\underline{\eta_1}-\eta_0$, где $\underline{\eta_1}=hf(x_1,\ \underline{y_1})$.

Найденные значения заносим в раздел I основного бланка (таблица 47).

Таблица 47

Схема вычисления начального отрезка методом последовательных сближений

№ сближения	i	x	y	Δy	η	$\Delta\eta$	$\Delta^2\eta$	$\Delta^3\eta$
I	0	x_0	$\underline{y_0}$	$\underline{\Delta y_0}$	$\underline{\eta_0}$	$\underline{\Delta\eta_0}$		
	1	x_1	$\underline{y_1}$		$\underline{\eta_1}$			
II	0	x_0	$\underline{\underline{y_0}}$	$\underline{\underline{\Delta y_0}}$	η_0	$\underline{\underline{\Delta\eta_0}}$	$\underline{\underline{\Delta^2\eta_0}}$	
	1	x_1	$\underline{\underline{y_1}}$	$\underline{\underline{\Delta y_1}}$	$\underline{\underline{\eta_1}}$	$\underline{\underline{\Delta\eta_1}}$		
	2	x_2	$\underline{\underline{y_2}}$		$\underline{\underline{\eta_2}}$			
III	0	x_0	$\underline{\underline{\underline{y_0}}}$	$\underline{\underline{\underline{\Delta y_0}}}$	η_0	$\underline{\underline{\underline{\Delta\eta_0}}}$	$\underline{\underline{\underline{\Delta^2\eta_0}}}$	$\underline{\underline{\underline{\Delta^3\eta_0}}}$
	1	x_1	$\underline{\underline{\underline{y_1}}}$	$\underline{\underline{\underline{\Delta y_1}}}$	$\underline{\underline{\underline{\eta_1}}}$	$\underline{\underline{\underline{\Delta\eta_1}}}$	$\underline{\underline{\underline{\Delta^2\eta_1}}}$	
	2	x_2	$\underline{\underline{\underline{y_2}}}$	$\underline{\underline{\underline{\Delta y_2}}}$	$\underline{\underline{\underline{\eta_2}}}$	$\underline{\underline{\underline{\Delta\eta_2}}}$		
	3	x_3	$\underline{\underline{\underline{y_3}}}$		$\underline{\underline{\underline{\eta_3}}}$			

Далее переходим ко второму сближению. Для этого, используя данные первого раздела, вычисляем приближенные значения Δy_0 и Δy_1 (эти значения, как и значения, связанные с ними, подчеркиваем двумя чертами) по двухчленным формулам:

$$\underline{\underline{\Delta y_0}} = \eta_0 + \frac{1}{2}\underline{\Delta\eta_0}, \tag{7}$$

$$\underline{\underline{\Delta y_1}} = \underline{\eta_1} + \frac{1}{2}\underline{\Delta\eta_0}. \tag{8}$$

Двухчленные формулы получаются соответственно из формулы (5) при $i=0$ и из формулы (2) при $i=1$ в результате откидывания разностей порядка выше первого.

Таким образом, получаем возможность найти

$$\underline{\underline{y_1}} = y_0 + \underline{\underline{\Delta y_0}} \quad \text{и} \quad \underline{\underline{y_2}} = \underline{\underline{y_1}} + \underline{\underline{\Delta y_1}},$$

в результате чего можно вычислить

$$\underline{\underline{\eta_1}} = hf(x_1, \underline{\underline{y_1}}), \quad \underline{\underline{\eta_2}} = hf(x_2, \underline{\underline{y_2}})$$

и составить разности

$$\underline{\underline{\Delta\eta_0}} = \underline{\underline{\eta_1}} - \eta_0, \quad \underline{\underline{\Delta\eta_1}} = \underline{\underline{\eta_2}} - \underline{\underline{\eta_1}}, \quad \underline{\underline{\Delta^2\eta_0}} = \underline{\underline{\Delta\eta_1}} - \underline{\underline{\Delta\eta_0}}.$$

Полученные результаты вписываем в раздел II основного бланка (таблица 47).

Для нахождения третьего сближения применяем трехчленные формулы, которые получаются из формулы (5) при $i=0$, из формулы (3) при $i=1$ и формулы (2) при $i=2$ после откидывания разностей третьего порядка. А именно, используя данные раздела II, вычисляем приближенные значения Δy_0, Δy_1, Δy_2 (подчеркнуты тремя линиями) по трехчленным формулам:

$$\underline{\underline{\underline{\Delta y_0}}} = \eta_0 + \frac{1}{2}\underline{\underline{\underline{\Delta\eta_0}}} - \frac{1}{12}\underline{\underline{\underline{\Delta^2\eta_0}}}, \tag{9}$$

$$\underline{\underline{\underline{\Delta y_1}}} = \underline{\underline{\eta_1}} + \frac{1}{2}\underline{\underline{\underline{\Delta\eta_1}}} - \frac{1}{12}\underline{\underline{\underline{\Delta^2\eta_0}}}, \tag{10}$$

$$\underline{\underline{\underline{\Delta y_2}}} = \underline{\underline{\eta_2}} + \frac{1}{2}\underline{\underline{\underline{\Delta\eta_1}}} + \frac{5}{12}\underline{\underline{\underline{\Delta^2\eta_0}}}. \tag{11}$$

Отсюда можно найти

$$\underline{\underline{y_1}} = y_0 + \underline{\underline{\underline{\Delta y_0}}}, \qquad \underline{\underline{y_2}} = \underline{\underline{y_1}} + \underline{\underline{\underline{\Delta y_1}}}, \qquad \underline{\underline{y_3}} = \underline{\underline{y_2}} + \underline{\underline{\underline{\Delta y_2}}}$$

и вычислить $\underline{\underline{\eta_1}}$, $\underline{\underline{\eta_2}}$, $\underline{\underline{\eta_3}}$. После этого можно заполнить раздел III основного бланка (таблица 47), найдя нужные разности обычным порядком.

Для контроля перевычисляем Δy_0, Δy_1 и Δy_2 по полным четырехчленным формулам (5) — при $i=0$, (4) — при $i=1$ и (3) — при $i=2$:

$$\Delta y_0 = \eta_0 + \frac{1}{2}\underline{\underline{\underline{\Delta\eta_0}}} - \frac{1}{12}\underline{\underline{\underline{\Delta^2\eta_0}}} + \frac{1}{24}\underline{\underline{\underline{\Delta^3\eta_0}}}, \tag{12}$$

$$\Delta y_1 = \underline{\underline{\eta_1}} + \frac{1}{2}\underline{\underline{\underline{\Delta\eta_1}}} - \frac{1}{12}\underline{\underline{\underline{\Delta^2\eta_1}}} + \frac{1}{24}\underline{\underline{\underline{\Delta^3\eta_0}}}, \tag{13}$$

$$\Delta y_2 = \underline{\underline{\eta_2}} + \frac{1}{2}\underline{\underline{\underline{\Delta\eta_2}}} - \frac{1}{12}\underline{\underline{\underline{\Delta^2\eta_1}}} - \frac{1}{24}\underline{\underline{\underline{\Delta^3\eta_0}}}. \tag{14}$$

Обычно перевычисление Δy_0 и Δy_1 по формулам (12) и (13) в пределах заданной точности не дает ничего нового по сравнению с $\underline{\underline{\underline{\Delta y_0}}}$ и $\underline{\underline{\underline{\Delta y_1}}}$. Что касается значения Δy_2, то, как правило, оно отлично от $\underline{\underline{\underline{\Delta y_2}}}$. В этом случае следует исправить $\underline{\underline{\underline{\Delta y_2}}}$ на Δy_2. Практически чаще всего приходится менять последний десятичный знак, и процедура сводится к тому, что этот знак берется в скобки и дописывается новый десятичный знак. Затем вносятся соответствующие исправления в значения $\underline{\underline{y_3}}$, $\underline{\underline{\eta_3}}$, $\underline{\underline{\underline{\Delta\eta_2}}}$, $\underline{\underline{\underline{\Delta^2\eta_1}}}$ и $\underline{\underline{\underline{\Delta^3\eta_0}}}$. После этого можно перейти к нормальному ходу вычислений по формулам Адамса.

Если $\underline{\underline{\Delta y_0}}$ и $\underline{\underline{\Delta y_1}}$ отличны соответственно от Δy_0 и Δy_1 в заданных десятичных знаках, то процесс сближений не может считаться законченным. В таком случае следует или найти четвертое сближение, полностью перевычислив с помощью формул (12)—(14) раздел III основного бланка, и затем для контроля проверить по тем же формулам устойчивость величин Δy_0, Δy_1, Δy_2, или же повторить весь ход вычислений с меньшим шагом.

Заметим, что при выборе шага h для метода сближений лучше ошибаться в меньшую сторону, чем в бо́льшую. Дело в том, что при малом шаге можно быстро закончить сближения, а затем, удвоив шаг, определить недостающие значения одним из регулярных методов, в то время как при завышенном шаге всю работу приходится переделывать заново.

Замечание. Можно указать более точный вариант метода последовательных сближений.

А именно, найдя $\underline{\underline{\Delta y_0}}$ по формуле (7), мы имеем возможность вычислить $\underline{\underline{y_1}}$, $\underline{\underline{\eta_1}}$ и $\underline{\underline{\Delta\eta_0}}$, после чего $\underline{\underline{\Delta y_1}}$ определяется по более точной формуле

$$\underline{\underline{\Delta y_1}} = \underline{\underline{\eta_1}} + \frac{1}{2}\underline{\underline{\Delta\eta_0}}.$$

Аналогичным образом можно поступить при вычислении $\underline{\underline{\Delta y_2}}$. А именно, вычислив $\underline{\underline{\Delta y_0}}$ и $\underline{\underline{\Delta y_1}}$ по формулам (9) и (10), находим $\underline{\underline{y_1}}$ и $\underline{\underline{y_2}}$ и вычисляем $\underline{\underline{\eta_1}}$, $\underline{\underline{\eta_2}}$, $\underline{\underline{\Delta\eta_0}}$, $\underline{\underline{\Delta\eta_1}}$, $\underline{\underline{\Delta^2\eta_0}}$. Затем $\underline{\underline{\Delta y_2}}$ определяется по формуле

$$\underline{\underline{\Delta y_2}} = \underline{\underline{\eta_2}} + \frac{1}{2}\underline{\underline{\Delta\eta_1}} + \frac{5}{12}\underline{\underline{\Delta^2\eta_0}}.$$

Пример 1. Методом последовательных сближений для интеграла дифференциального уравнения

$$y' = x + y, \quad y(0) = 1 \tag{15}$$

найти значения $y_1 = y(0{,}1)$, $y_2 = y(0{,}2)$, $y_3 = y(0{,}3)$, приняв шаг $h = 0{,}1$.

Решение. Вычисление будем вести с тремя десятичными знаками после запятой.

Полагая $\eta = 0{,}1\,(x+y)$ и применяя формулу (6), получаем

$$\underline{\Delta y_0} = 0{,}100, \qquad \underline{y_1} = 1{,}000 + 0{,}100 = 1{,}100.$$

После этого заполняем раздел I основного бланка (таблица 48), далее с помощью двухчленных формул (7) и (8) находим

$$\underline{\underline{\Delta y_0}} = 0{,}100 + \frac{1}{2}\,0{,}020 = 0{,}110, \qquad \underline{\underline{\Delta y_1}} = 0{,}120 + \frac{1}{2}\,0{,}020 = 0{,}130,$$

что дает возможность заполнить раздел II основного бланка (таблица 48). Наконец, используя трехчленные формулы (9) — (11), будем иметь:

$$\underline{\underline{\underline{\Delta y_0}}} = 0{,}100 + \frac{1}{2}\cdot 0{,}021 - \frac{1}{12}\cdot 0{,}002 = 0{,}110,$$

$$\underline{\underline{\underline{\Delta y_1}}} = 0{,}121 + \frac{1}{2}\cdot 0{,}023 - \frac{1}{12}\cdot 0{,}002 = 0{,}132,$$

$$\underline{\underline{\underline{\Delta y_2}}} = 0{,}144 + \frac{1}{2}\cdot 0{,}023 + \frac{5}{12}\cdot 0{,}002 = 0{,}156.$$

С помощью полученных данных заполняем раздел III основного бланка (таблица 48).

Таблица 48

Вычисление начального отрезка интеграла дифференциального уравнения (15) методом последовательных сближений

	i	x	y	Δy	η	$\Delta\eta$	$\Delta^2\eta$	$\Delta^3\eta$
I	0	0	1,000	0,100	0,100	20		
	1	0,1	1,100		0,120			
II	0	0	1,000	0,110	0,100	21	2	
	1	0,1	1,110	0,130	0,121	23		
	2	0,2	1,240		0,144			
III	0	0	1,000	0,110	0,100	21	2	1
	1	0,1	1,110	0,132	0,121	23	3	
	2	0,2	1,242	0,15(6)7	0,144	26		
	3	0,3	1,39(8)9		0,170			

Для контроля производим пересчет Δy_0, Δy_1 и Δy_2 по четырехчленным формулам (12)—(14). В результате приходится исправлять на единицу последнего разряда лишь значение Δy_2, помещая неточную цифру в скобки.

Аналогичным способом в случае необходимости исправляются и значения y_3, η_3, $\Delta\eta_2$, $\Delta^2\eta_1$ и $\Delta^3\eta_0$. В данном случае с точностью до третьего десятичного знака эти величины остаются неизменными.

Таким образом, окончательно имеем $y(0,\ 1) = 1{,}110$, $y(0,\ 2) = 1{,}242$, $y(0,\ 3) = 1{,}399$.

Для сравнения приводим точные значения (см. § 7):

$$y(0,\ 1) = 1{,}1103\ldots;\quad y(0,\ 2) = 1{,}2428\ldots;\quad y(0,\ 3) = 1{,}3997\ldots$$

§ 10. Метод Милна

Одним из наиболее простых и практически удобных методов численного интегрирования дифференциальных уравнений является метод Милна (W. E. Milne) [19], [21]. Дадим сначала описание этого метода, а затем остановимся на выводе необходимых формул, применяемых в методе Милна.

Пусть дано дифференциальное уравнение

$$y' = f(x, y) \tag{1}$$

с начальным условием

$$y(x_0) = y_0. \tag{2}$$

Выбрав шаг h, положим, как обычно,

$$x_i = x_0 + ih, \quad y_i = y(x_i), \quad y'_i = f(x_i, y_i) \quad (i = 0, 1, 2, \ldots).$$

Первые четыре значения искомого решения y_0, y_1, y_2, y_3 («начальный отрезок») находим, используя начальное условие (2) и применяя какой-либо метод, описанный выше, например метод последовательных приближений или метод Рунге—Кутта. Тем самым будут известны y'_i $(i = 0, 1, 2, 3)$.

Дальнейшие значения $y_i = y(x_i)$ $(i = 4, 5, \ldots)$ последовательно определяются по следующей схеме: предполагая, что y_{i-1}, y_{i-2}, y_{i-3}, y_{i-4} известны,

1) вычисляем первое приближение $y_i^{(1)}$ для ближайшего следующего значения y_i по формуле

$$y_i^{(1)} = y_{i-4} + \frac{4h}{3}(2y'_{i-3} - y'_{i-2} + 2y'_{i-1}) \quad (i = 4, 5, \ldots); \tag{3}$$

2) значение $y_i^{(1)}$ подставляем в дифференциальное уравнение (1) и определяем соответствующее значение $y_i'^{(1)} = f(x_i, y_i^{(1)})$;

3) находим второе приближение $y_i^{(2)}$ по формуле

$$y_i^{(2)} = y_{i-2} + \frac{h}{3}(y'_{i-2} + 4y'_{i-1} + y_i'^{(1)}) \quad (i = 4, 5, \ldots). \tag{4}$$

Милн показал, что абсолютная погрешность значения $y_i^{(2)}$ приближенно равна

$$\varepsilon_i = \frac{1}{29}\left| y_i^{(2)} - y_i^{(1)} \right|. \tag{5}$$

Поэтому, если $\varepsilon_i \leqslant \varepsilon$, где ε — заданная предельная погрешность решения, то можно положить $y_i \approx y_i^{(2)}$ и $y'_i \approx f(x_i, y_i^{(2)})$. В частности, это имеет место, если $y_i^{(1)}$ и $y_i^{(2)}$ совпадают в интересующих нас десятичных знаках.

Далее переходим к вычислению ближайшего следующего значения y_{i+1}, повторяя указанный выше процесс. В противном случае, если точность ε не обеспечена, следует уменьшить шаг h (начиная с известного места). При этом мы встретимся с неприятной необходимостью пересчета соответствующего «начального отрезка».

Таким образом, мы видим, что метод Милна выгодно отличается от других методов (например, от метода Рунге—Кутта) тем, что в нем производится корректирование каждого вновь полученного

частного значения интеграла уравнения без пересчета с измененным шагом.

Приступим теперь к выводу формул Милна (3)—(5). Для этого воспользуемся первой интерполяционной формулой Ньютона [23], написанной для производной y', в подходяще выбранной точке x_k, причем ограничимся разностями третьего порядка, что равносильно тому, что интеграл $y=y(x)$ дифференциального уравнения (1) аппроксимируется полиномом четвертой степени. Имеем

$$y'=y'_k+q\Delta y'_k+\frac{q(q-1)}{2!}\Delta^2 y'_k+\frac{q(q-1)(q-2)}{3!}\Delta^3 y'_k$$

или

$$y'=y'_k+q\Delta y'_k+\frac{1}{2}(q^2-q)\Delta^2 y'_k+\frac{1}{6}(q^3-3q^2+2q)\Delta^3 y'_k, \qquad (6)$$

где

$$q=\frac{x-x_k}{h}. \qquad (7)$$

Полагая $k=i-4$ в формуле (6) и почленно интегрируя полученную

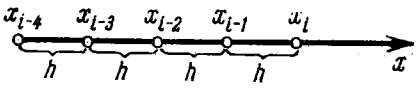

Рис. 32.

формулу по x в пределах от x_{i-4} до x_i (рис. 32), будем иметь

$$\int_{x_{i-4}}^{x_i} y'\,dx=\int_{x_{i-4}}^{x_i}\left[y'_{i-4}+q\Delta y'_{i-4}+\frac{q^2-q}{2}\Delta^2 y'_{i-4}+\right.$$
$$\left.+\frac{q^3-3q^2+2q}{6}\Delta^3 y'_{i-4}\right]dx.$$

Отсюда, учитывая, что $q=\frac{x-x_{i-4}}{h}$ и $dx=hdq$, находим

$$y_i-y_{i-4}=h\left\{y'_{i-4}\int_0^4 dq+\Delta y'_{i-4}\int_0^4 q\,dq+\right.$$
$$\left.+\Delta^2 y'_{i-4}\int_0^4\frac{q^2-q}{2}dq+\Delta^3 y'_{i-4}\int_0^4\frac{q^3-3q^2+2q}{6}dq\right\}=$$
$$=h\left(4y'_{i-4}+8\Delta y'_{i-4}+\frac{20}{3}\Delta^2 y'_{i-4}+\frac{8}{3}\Delta^3 y'_{i-4}\right). \qquad (8)$$

Так как

$$\Delta y'_{i-4} = y'_{i-3} - y'_{i-4},$$
$$\Delta^2 y'_{i-4} = y'_{i-2} - 2y'_{i-3} + y'_{i-4},$$
$$\Delta^3 y'_{i-4} = y'_{i-1} - 3y'_{i-2} + 3y'_{i-3} - y'_{i-4},$$

то, подставляя эти значения в формулу (8), после обычных упрощений получим первую формулу Милна:

$$y_i = y_{i-4} + \frac{4h}{3}(2y'_{i-3} - y'_{i-2} + 2y'_{i-1}). \tag{9}$$

Для вывода второй формулы Милна положим $k = i - 2$ в формуле (6) и проинтегрируем обе части получившегося выражения по x в пределах от x_{i-2} до x_i. Тогда, учитывая, что $q = \frac{x - x_{i-2}}{h}$ и $dq = \frac{dx}{h}$, будем иметь

$$\int_{x_{i-2}}^{x_i} y'\,dx = h\int_0^2 \left[y'_{i-2} + q\,\Delta y'_{i-2} + \frac{1}{2}(q^2 - q)\,\Delta^2 y'_{i-2} + \right.$$
$$\left. + \frac{1}{6}(q^3 - 3q^2 + 2q)\,\Delta^3 y'_{i-2}\right] dq.$$

Отсюда, выполняя квадратуры, получим

$$y_i - y_{i-2} = h\left(2y'_{i-2} + 2\Delta y'_{i-2} + \frac{1}{3}\Delta^2 y'_{i-2}\right). \tag{10}$$

Подставляя в формулу (10) известные выражения

$$\Delta y'_{i-2} = y'_{i-1} - y'_{i-2}, \quad \Delta^2 y'_{i-2} = y'_i - 2y'_{i-1} + y'_{i-2},$$

придем ко второй формуле Милна:

$$y_i = y_{i-2} + \frac{h}{3}(y'_{i-2} + 4y'_{i-1} + y'_i). \tag{11}$$

Заметим, что для дифференциального уравнения

$$y' = f(x), \quad y(x_0) = y_0,$$

правая часть которого не зависит от искомой функции y, формула (11) идентична с формулой Симпсона для интеграла

$$y_i - y_{i-2} = \int_{x_{i-2}}^{x_i} f(x)\,dx.$$

Для вывода контрольной формулы (5) для погрешности ε_i второго приближения $y_i^{(2)}$ оценим главные члены погрешностей $\varepsilon_i^{(1)}$ и $\varepsilon_i^{(2)}$ первой и второй формул Милна. Учитывая отброшенные в интерполя-

ционной формуле Ньютона (6) разности четвертого порядка, с точностью до разностей пятого порядка будем иметь

$$\left.\begin{aligned} \varepsilon_i^{(1)} &\approx h\int_0^4 \frac{1}{24}(q^4-6q^3+11q^2-6q)\,\Delta^4 y'_{i-4}\,dq=\frac{28}{90}h\Delta^4 y'_{i-4},\\ \varepsilon_i^{(2)} &\approx h\int_0^2 \frac{1}{24}(q^4-6q^3+11q^2-6q)\,\Delta^4 y'_{i-2}\,dq=\\ &\qquad\qquad =-\frac{h}{90}\Delta^4 y'_{i-2}. \end{aligned}\right\} \quad (12)$$

Отсюда, считая, что четвертая разность $\Delta^4 y'_i$ постоянна на интервале длины $4h$, получим $\varepsilon_i^{(1)}=-28\varepsilon_i^{(2)}$.

Так как, очевидно, $y_i=y_i^{(1)}+\varepsilon_i^{(1)}=y_i^{(1)}-28\,\varepsilon_i^{(2)}$ и $y_i=y_i^{(2)}+\varepsilon_i^{(2)}$, то

$$\varepsilon_i^{(2)}=\frac{1}{29}(y_i^{(1)}-y_i^{(2)}).$$

Следовательно, имеем контрольную формулу Милна:

$$\varepsilon_i=|\varepsilon_i^{(2)}|=\frac{|y_i^{(1)}-y_i^{(2)}|}{29}. \quad (13)$$

Заметим, что если шаг h достаточно мал, то приближенно можно положить

$$\frac{\Delta^4 y'_i}{h^4}\approx y^{\mathrm{V}}(x_i).$$

Пусть $a\leqslant x\leqslant b$ — интересующий нас отрезок, на котором строится решение дифференциального уравнения (1), и $h=\frac{b-a}{n}$. Тогда из формул (12) следует, что предельная абсолютная погрешность на отрезке $[a, b]$ приближенного решения $y_i=y(x_i)$ выражается следующим образом:

$$E=\frac{h^5}{90}M_5 n=\frac{b-a}{90}M_5 h^4,$$

где $M_5=\max|y^{\mathrm{V}}(x)|$ при $a\leqslant x\leqslant b$.

Таким образом, суммарная ошибка метода Милна есть величина порядка h^4.

Пример 1. Дано дифференциальное уравнение

$$y'=x+y \quad (14)$$

с начальным условием $y(0)=1$. Методом Милна вычислить $y(0{,}5)$ с точностью до 0,001.

Решение. Задаемся $h=0{,}1$; так как погрешность результата получается порядка $h^4=0{,}0001$, то заданная точность практически достигается. Вычисления ведем с одним запасным знаком.

Значения y_0 и y_0' нам известны непосредственно из начального условия и уравнения (14): $y_0=1$, $y_0'=1$. Остальные первые три значения

$$y_i=y(x_i) \qquad (i=1,\ 2,\ 3)$$

находим каким-либо другим способом численного интегрирования дифференциального уравнения.

Из уравнения (14) получаем соответствующие значения

$$y_i'=y'(x_i) \qquad (i=1,\ 2,\ 3).$$

В таблице 49 приведены эти значения, найденные методом Рунге—Кутта (см. § 7, табл. 40).

Таблица 49

Начальный отрезок для метода Милна

i	x_i	y_i	y_i'
0	0	1	1
1	0,1	1,1103	1,2103
2	0,2	1,2427	1,4427
3	0,3	1,3996	1,6996

Займемся вычислением значений y_4 и y_5 по формулам Милна.

1) Вычисление y_4. Применяя первую формулу Милна (3), получим

$$y_4^{(1)}=1+\frac{4\cdot 0{,}1}{3}(2\cdot 1{,}2103-1{,}4427+2\cdot 1{,}6996)=1{,}5836.$$

Отсюда, подставив в дифференциальное уравнение, будем иметь $y_4'^{(1)}=1{,}9836$. Далее, применяя вторую формулу Милна (4), находим

$$y_4^{(2)}=1{,}2427+\frac{0{,}1}{3}(1{,}4427+4\cdot 1{,}6996+1{,}9836)=1{,}5835.$$

Таким образом, найденные значения $y_4^{(1)}$ и $y_4^{(2)}$ совпадают в сохраняемых нами трех десятичных знаках, поэтому считаем $y_4=y_4^{(2)}=1{,}5835$. Отсюда

$$y_4'=y_4'^{(2)}=\underline{1{,}9835.}$$

2) Вычисление y_5. По первой формуле Милна имеем

$$y_5^{(1)}=1{,}1103+\frac{4\cdot 0{,}1}{3}(2\cdot 1{,}4427-1{,}6996+2\cdot 1{,}9835)=1{,}7973;$$

соответственно

$$y_5'^{(1)}=0{,}5+1{,}973=2{,}2973.$$

По второй формуле Милна будем иметь

$$y_5^{(2)} = 1,3996 + \frac{0,1}{3}(1,6996 + 4 \cdot 1,9835 + 2,2973) = 1,7973.$$

Так как $y_5^{(1)}$ и $y_5^{(2)}$ совпадают, то можно положить $y_5 = 1,7973$. Следовательно, $y(0,5) = 1,797$.

Отметим, что благодаря малости шага h здесь нет необходимости оценивать погрешность приближенных значений искомого решения.

Интегрирование дифференциального уравнения по методу Милна удобно производить по следующей схеме (таблицы 50, 51).

Таблица 50

Основная таблица для метода Милна

i	x_i	$y_i^{(1)}$	$y_i^{(2)} = y_i$	ε_i	$y_i'^{(1)}$	$y_i'^{(2)} = y_i'$
0	0		1			1
1	0,1		1,1103			1,2103
2	0,2		1,2427			1,4427
3	0,3		1,3996			1,6996
4	0,4	1,5836	1,5835	0	1,9836	1,9835
5	0,5	1,7973	1,7973	0	2,2973	2,2973

Метод Милна может быть применен также к интегрированию систем дифференциальных уравнений.

Пусть дана система дифференциальных уравнений

$$\frac{dy}{dx} = f(x, y) \tag{15}$$

с начальными условиями

$$y(x_0) = y_0. \tag{16}$$

Выберем шаг h и положим

$$x_i = x_0 + ih, \quad y_i = y(x_i) \qquad (i = 0, 1, 2, \ldots),$$

где $y = y(x)$ — искомая вектор-функция.

Каким-либо методом определим дополнительные начальные значения y_1, y_2, y_3.

Из уравнения (15) найдем соответствующие значения производной

$$y_i' = f(x_i, y_i) \qquad (i = 0, 1, 2, 3).$$

Тогда дальнейшие значения y_i $(i = 4, 5, 6, \ldots)$ искомой вектор-функции приближенно шаг за шагом могут быть вычислены по следующим формулам:

$$\text{I. } y_i^{(1)} = y_{i-4} + \frac{4h}{3}(2y_{i-3}' - y_{i-2}' + 2y_{i-1}').$$

$$\text{II. } y_i^{(2)} = y_{i-2} + \frac{h}{3}(y_{i-2}' + 4y_{i-4}' + y_i'^{(1)}),$$

где $y_i'^{(1)} = f(x_i, y_i^{(1)})$.

Таблица 51

Вспомогательная таблица для метода Милна

i	4	5
$2y'_{i-3}$ $-y'_{i-2}$ $2y'_{i-1}$	2,4206 −1,4427 3,3992	2,8854 −1,6996 3,9670
$\sum_i^{(1)}$	4,3771	5,1528
$\frac{4}{3}h\sum_i^{(1)}$ y_{i-4}	0,5836 1	0,6870 1,1103
$y_i^{(1)}$	1,5836	1,7973
y'_{i-2} $4y'_{i-1}$ $y_i'^{(1)}$	1,4427 6,7984 1,9836	1,6996 7,9340 2,2973
$\sum_i^{(2)}$	10,2247	11,9309
$\frac{h}{3}\sum_i^{(2)}$ y_{i-2}	0,3408 1,2427	0,3977 1,3996
$y_i^{(2)}$	1,5835	1,7973

III. Если

$$\varepsilon_i = \frac{\|y_i^{(1)} - y_i^{(2)}\|}{29} \leqslant \varepsilon \text{ *)},$$

где ε — заданная предельная погрешность, то полагаем $y_i \approx y_i^{(2)}$.

---

*) Под *нормой* вектора $y = (y_1, y_2, \ldots, y_n)$ здесь понимается число

$$\|y\| = \max_j |y_j|.$$

Пример 2. Методом Милна найти на отрезке [0, 1] интеграл системы дифференциальных уравнений

$$\frac{dy}{dx}=\frac{z}{x}, \quad \frac{dz}{dx}=-xy, \tag{17}$$

удовлетворяющий начальным условиям $y(0)=1$, $z(0)=0$.

Решение. Для удобства записи положим

$$f=\begin{bmatrix} f_1 \\ f_2 \end{bmatrix},$$

где $f_1=\frac{z}{x}$ и $f_2=-xy$.

Заметим, что система (17) обладает особенностью при $x=0$. Вследствие этого $y'(0)$ имеет форму неопределенности, которую раскрываем по правилу Лопиталя:

$$y'(0)=\lim_{x\to 0}\frac{z}{x}=\lim_{x\to 0}\frac{z'}{1}=z'(0)=0.$$

Выберем шаг $h=0{,}2$. Начальный отрезок для метода Милна вычисляем методом Рунге—Кутта с пятью десятичными знаками после запятой. Схема вычислений дается в таблице 52.

Имея начальные значения $Y_i=(y_i, z_i)$ $(i=0, 1, 2, 3)$, дальнейшие вычисления проводим методом Милна, используя формулы I, II, III.

Результаты вычислений даны в таблицах 53 и 54.

Таким образом, мы получаем

$$y(1)=0{,}76520; \quad z(1)=-0{,}44000.$$

По вычислениям методом Адамса (см. [20]) имеем

$$y(1)=0{,}76520; \quad z(1)=-0{,}44005.$$

Заметим, что в нашем случае известно точное решение системы (17), а именно:

$$y=J_0(x), \quad z=-xJ_1(x),$$

где J_0 и J_1—функции Бесселя нулевого и первого индекса соответственно. По таблицам [17] имеем

$$y(1)=J_0(1)=0{,}765198\ldots, \quad z(1)=-J_1(1)=-0{,}440051\ldots$$

Пример 3. Методом Милна найти на отрезке $0\leqslant t\leqslant 0{,}5$ частное решение системы

$$\left.\begin{aligned} x' &= -2x+5z, \\ y' &= -(1-\sin t)x-y+3z, \\ z' &= -x+2z, \end{aligned}\right\} \tag{18}$$

удовлетворяющее начальным условиям

$$x(0)=x_0=2, \quad y(0)=y_0=1, \quad z(0)=z_0=1.$$

Шаг $h=\Delta t$ принять равным 0,1.

Таблица 52

Вычисление начального отрезка для метода Милна в случае системы дифференциальных уравнений

i	x	y	z	$f_1=\frac{z}{x}$	$f_2=-xy$	$k^{(1)}=0{,}2f_1$	$k^{(2)}=0{,}2f_2$	Δy	Δz
0	0	1	0	0	0	0	0	0	0
	0,1	1	0	0	—0,10000	0,00000	—0,02000	0,00000	—0,04000
	0,1	1	—0,01000	—0,10000	—0,10000	—0,02000	—0,02000	—0,04000	—0,04000
	0,2	0,98000	—0,02000	—0,10000	—0,19600	—0,02000	—0,03920	—0,02000	—0,03920
								$-0{,}06000\cdot\frac{1}{6}=$ $=-0{,}01000$	$-0{,}11920\cdot\frac{1}{6}=$ $=-0{,}01897$
1	0,2	0,99000	—0,01987	—0,09933	—0,19800	—0,01987	—0,03960	—0,01987	—0,03960
	0,3	0,98007	—0,03967	—0,13222	—0,29402	—0,02644	—0,05880	—0,05888	—0,11760
	0,3	0,97678	—0,04927	—0,16423	—0,29303	—0,03285	—0,05861	—0,06570	—0,11722
	0,4	0,95715	—0,07848	—0,19620	—0,38286	—0,03924	—0,07657	—0,03924	—0,07657
								$-0{,}17769\cdot\frac{1}{6}=$ $=-0{,}02962$	$-0{,}35099\cdot\frac{1}{6}=$ $=-0{,}05850$
2	0,4	0,96038	—0,07837	—0,19592	—0,38416	—0,3918	—0,07683	—0,03928	—0,07683
	0,5	0,94080	—0,11679	—0,23358	—0,47040	—0,04672	—0,09408	—0,09344	—0,18816
	0,5	0,93703	—0,12541	—0,25082	—0,46852	—0,05016	—0,09370	—0,10032	—0,18740
	0,6	0,91027	—0,17207	—0,28678	—0,54614	—0,05736	—0,10923	—0,05736	—0,10923
								$-0{,}29030\cdot\frac{1}{6}=$ $=-0{,}04838$	$-0{,}56162\cdot\frac{1}{6}=$ $=-0{,}09360$
3	0,6	0,91200	—0,17197	—0,28662	—0,54721				

Таблица 53

Основная таблица для метода Милна в случае системы дифференциальных уравнений

i	x_i	$Y_i^{(1)}$		$Y_i'^{(1)}$		$Y_i^{(2)}=Y_i$		$Y_i'^{(2)}=Y_i'$		$\bar{\varepsilon}_i$	
		$y_i^{(1)}$	$z_i^{(1)}$	$y_i'^{(1)}$	$z_i'^{(1)}$	$y_i^{(2)}$	$z_i^{(2)}$	$y_i'^{(2)}=y_i'$	$z_i'^{(2)}=z_i'$	$\varepsilon_{i,1}$	$\varepsilon_{i,2}$
0	0					1	0	0	0		
1	0,2					0,99000	—0,01987	—0,09933	—0,19800		
2	0,4					0,96038	—0,07837	—0,19592	—0,38416		
3	0,6					0,91200	—0,17197	—0,28662	—0,54721		
4	0,8	0,84737	—0,29500	—0,36875	—0,67790	0,84630	—0,29510	—0,36887	—0,67704	$4\cdot10^{-5}$	$\frac{1}{3}\cdot10^{-5}$
5	1,0	0,76522	—0,43987	—0,43987	—0,76522	0,76520	—0,44000			$2\cdot10^{-5}$	$\frac{1}{2}\cdot10^{-5}$

Таблица 54

Вспомогательная таблица для метода Милна в случае системы дифференциальных уравнений

i	4		5	
	y_i	z_i	y_i	z_i
$2Y'_{i-3}$	−0,19866	−0,39600	−0,39184	−0,76832
$-Y'_{i-2}$	0,19592	0,38416	0,28862	0,54721
$2Y'_{i-1}$	−0,57324	−0,09442	−0,73774	−1,35408
$\sum_i^{(1)}$	−0,57238	−1,10626	−0,84296	−1,57519
$\frac{4}{3}h\sum_i^{(1)}$	−0,15263	−0,29500	−0,22478	−0,42000
Y_{i-4}	1,00000	0,00000	0,99000	−0,01987
$Y_i^{(1)}$	0,84737	−0,29500	−0,76522	−0,43987
Y'_{i-2}	−0,19592	−0,38416	−0,28662	−0,54721
$4Y'_{i-1}$	−1,14648	−2,18884	−1,47548	−2,70808
$Y_i'^{(1)}$	−0,36875	−0,67790	−0,43987	−0,76522
$\sum_i^{(2)}$	−1,71115	−3,25090	−2,20197	−4,02051
$\frac{h}{3}\sum_i^{(2)}$	−0,11408	−0,21673	−0,14680	−0,26803
Y_{i-2}	0,96038	−0,07837	0,91200	−0,17197
$Y_i^{(2)}$	0,84630	−0,29510	0,76520	−0,44000

Решение. Начальный отрезок для искомого решения вычисляем методом Рунге—Кутта. Для краткости положим

$$\boldsymbol{X}=\begin{bmatrix} x \\ y \\ z \end{bmatrix}, \quad \boldsymbol{k}=h\boldsymbol{X}'=0{,}1\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix}$$

и $\Delta\boldsymbol{X}=\frac{1}{6}\sum$, $\sum=q_1\boldsymbol{k}_1+q_2\boldsymbol{k}_2+q_3\boldsymbol{k}_3+q_4\boldsymbol{k}_4$, где $q_1=q_4=1$, $q_2=q_3=2$.

Результаты вычислений приведены в табл. 55.

Таблица 55

Нахождение начального отрезка решения X методом Рунге — Кутта

i	t	X	$k=0,1X'$	qk	$\Delta X=\frac{1}{6}\sum$
0	0	2 1 1	0,1 0 0	0,1 0 0	0,08984 0,00497 —0,00500
	0,05	2,05 1 1	0,09 0,00525 —0,005	0,18 0,01050 —0,010	
	0,05	2,045 1,00262 0,99750	0,08975 0,00471 —0,00500	0,17950 0,00942 —0,01000	
	0,1	2,08975 1,00471 0,99500	0,07955 0,00992 —0,00998	0,07955 0,00992 —0,00998	
1	0,1	2,08984 1,00497 0,99500	0,07953 0,00988 —0,00998	0,07953 0,00988 —0,00998	0,06896 0,01456 —0,01494
	0,15	2,12960 1,00991 0,99001	0,06908 0,01480 —0,01496	0,13816 0,02960 —0,02992	
	0,15	2,12438 1,01237 0,98752	0,06888 0,01433 —0,01493	0,13777 0,02866 —0,02986	
	0,2	2,15872 1,01930 0,98007	0,05829 0,01911 —0,01986	0,05829 0,01911 —0,01986	
2	0,2	2,15880 1,01953 0,98006	0,05827 0,01907 —0,01987	0,05827 0,01907 —0,01987	0,04739 0,02313 —0,02473
	0,25	2,18794 1,02907 0,97012	0,04747 0,02346 —0,02477	0,09494 0,04692 —0,04954	
	0,25	2,18254 1,03126 0,96758	0,04733 0,02292 —0,02472	0,09466 0,04584 —0,04944	
	0,3	2,20613 1,04245 0,95534	0,03644 0,02694 0,02954	0,03644 0,02694 0,02954	

Дальнейшие вычисления произведены по формулам I, II и III. Полученные результаты помещены в таблице 56, причем промежуточные выкладки опущены. Для сравнения в последней графе таблицы 56 приведены компоненты точного решения $\tilde{X}$ с пятью верными десятичными знаками после запятой.

Таблица 56

Вычисление решения X методом Милна

i	t	$X=X^{(2)}$	$X'=X^{(2)'}$	$X^{(1)}$	$X^{(1)'}$	$\varepsilon=\frac{\lvert X^{(1)}-X^{(2)}\rvert}{29}$	$\tilde{X}$
0	0	2 1 1	1 0 0				2 1 1
1	0,1	2,08984 1,00497 0,99500	0,79532 0,09882 —0,09984				2,08983 1,00497 0,99500
2	0,2	2,15880 1,01953 0,98006	0,58270 0,19074 —0,19868				2,15881 1,01954 0,98007
3	0,3	2,20619 1,04266 0,95533	0,36427 0,26911 —0,29553				2,20620 1,04267 0,95534
4	0,4	2,23153 1,07270 0,92106	0,14224 0,32795 0,38941	2,23152 1,07268 0,92106	0,36427 0,32798 0,38940	0 0 0	2,23154 1,07271 0,92106
5	0,5	2,23459 1,10742 0,87758		2,23458 1,10740 0,87758	—0,08126 0,36209 —0,47942	0 0 0	2,23459 1,10743 0,87758

§ 11. Методы, основанные на применении производных высших порядков

До сих пор для численного интегрирования дифференциального уравнения первого порядка

$$y' = f(x, y) \tag{1}$$

с начальным условием

$$y(x_0) = y_0 \tag{2}$$

мы применяли формулы, в которых явно используется лишь первая производная $y_i' = y_i'(x_i)$ искомого решения.

Однако если использовать формулы, явно содержащие производные высших порядков от искомого решения, то можно указать методы, дающие более точный результат на данном промежутке без увеличения числа шагов.

Выведем соответствующие формулы, предполагая, что правая часть уравнения (1) дифференцируема достаточное число раз.

Пусть y_i, y'_i, y''_i — значения искомого решения $y=y(x)$ и, соответственно, значения его производных первого и второго порядков в точках $x_i=x_0+ih$ $(i=0, 1, 2, \ldots)$. Разлагая величины

$$y_{i+1}=y(x_i+h),\quad y'_{i+1}=y'(x_i+h),\quad y''_{i+1}=y''(x_i+h)$$

в ряды по степеням h, находим:

$$y_{i+1}=y_i+y'_ih+\frac{y''_i}{2!}h^2+\frac{y'''_i}{3!}h^3+\frac{y_i^{\mathrm{IV}}}{4!}h^4+\frac{y_i^{\mathrm{V}}}{5!}h^5+\ldots,$$

$$hy'_{i+1}=y'_ih+y''_ih^2+\frac{y'''_i}{2!}h^3+\frac{y_i^{\mathrm{IV}}}{3!}h^4+\frac{y_i^{\mathrm{V}}}{4!}h^5+\ldots,$$

$$h^2y''_{i+1}=y''_ih^2+y'''_ih^3+\frac{y_i^{\mathrm{IV}}}{2!}h^4+\frac{y_i^{\mathrm{V}}}{3!}h^5+\ldots$$

Из полученных формул исключим члены, содержащие y'''_i и y_i^{IV}. Для этого вторую формулу умножим на $-\frac{1}{2}$, а третью — на $\frac{1}{12}$ и сложим с первой. Будем иметь:

$$y_{i+1}-\frac{h}{2}y'_{i+1}+\frac{h^2}{12}y''_{i+1}=y_i+y'_i\left(1-\frac{1}{2}\right)h+$$
$$+y''_i\left(\frac{1}{2}-\frac{1}{2}+\frac{1}{12}\right)h^2+y'''_i\left(\frac{1}{6}-\frac{1}{4}+\frac{1}{12}\right)h^3+$$
$$+y_i^{\mathrm{IV}}\left(\frac{1}{24}-\frac{1}{12}+\frac{1}{24}\right)h^4+y_i^{\mathrm{V}}\left(\frac{1}{120}-\frac{1}{48}+\frac{1}{72}\right)h^5+\ldots$$

или

$$y_{i+1}=y_i+\frac{h}{2}(y'_i+y'_{i+1})+\frac{h^2}{12}(y''_i-y''_{i+1})+\frac{h^5}{720}y_i^{\mathrm{V}}+\ldots$$

Таким образом, с точностью до h^5 имеем приближенную формулу [14]

$$y_{i+1}=y_i+\frac{h}{2}(y'_i+y'_{i+1})+\frac{h^2}{12}(y''_i-y''_{i+1}). \tag{3}$$

Можно показать, что остаточный член формулы (3) равен

$$R_1=\frac{y^{\mathrm{V}}(\xi)}{720}h^5,$$

где $x_i<\xi<x_{i+1}$. Аналогично имеем:

$$y_{i+1}-y_{i-2}=y(x_i+h)-y(x_i-2h)=$$
$$=\left(y_i+hy_i'+\frac{h^2}{2}y_i''+\frac{h^3}{6}y_i'''+\frac{h^4}{24}y_i^{\mathrm{IV}}+\frac{h^5}{120}y_i^{\mathrm{V}}+\ldots\right)-$$
$$-\left(y_i-2hy_i'+\frac{4h^2}{2}y_i''-\frac{8h^3}{6}y_i'''+\frac{16h^4}{24}y_i^{\mathrm{IV}}-\frac{32h^5}{120}y_i^{\mathrm{V}}+\ldots\right)=$$
$$=3hy_i'-\frac{3h^2}{2}y_i''+\frac{3h^3}{2}y_i'''-\frac{5h^4}{8}y_i^{\mathrm{IV}}+\frac{11h^5}{40}y_i^{\mathrm{V}}+\ldots$$

и

$$y_i-y_{i-1}=y_i-y(x_i-h)=$$
$$=y_i-\left(y_i-hy_i'+\frac{h^2}{2}y_i''-\frac{h^3}{6}y_i'''+\frac{h^4}{24}y_i^{\mathrm{IV}}-\frac{h^5}{120}y_i^{\mathrm{V}}+\ldots\right)=$$
$$=hy_i'-\frac{h^2}{2}y_i''+\frac{h^3}{6}y_i'''-\frac{h^4}{24}y_i^{\mathrm{IV}}+\frac{h^5}{120}y_i^{\mathrm{V}}-\ldots$$

Отсюда

$$y_{i+1}-y_{i-2}-3(y_i-y_{i-1})=h^3y_i'''-\frac{h^4}{2}y_i^{\mathrm{IV}}+\frac{h^5}{4}y_i^{\mathrm{V}}+\ldots$$

С другой стороны,

$$y_i''-y_{i-1}''=y_i''-y_i''(x_i-h)=y_i''-\left(y_i''-hy_i'''+\frac{h^2}{2}y_i^{\mathrm{IV}}-\right.$$
$$\left.-\frac{h^3}{6}y_i^{\mathrm{V}}+\ldots\right)=hy_i'''-\frac{h^2}{2}y_i^{\mathrm{IV}}+\frac{h^3}{6}y_i^{\mathrm{V}}-\ldots$$

Поэтому

$$y_{i+1}-y_{i-2}-3(y_i-y_{i-1})=h^2(y_i''-y_{i-1}'')+\frac{h^5}{12}y_i^{\mathrm{V}}+\ldots$$

Таким образом, с точностью до h^5 имеем приближенную формулу

$$y_{i+1}=y_{i-2}+3(y_i-y_{i-1})+h^2(y_i''-y_{i-1}''). \tag{4}$$

Можно доказать, что остаточный член формулы (4) есть

$$R_2=\frac{h^5}{12}y^{\mathrm{V}}(\xi'),$$

где $x_{i-2}<\xi'<x_{i+1}$.

К формулам (3) и (4) присоединим выражения для производных y' и y'', а именно:

$$y'=f(x,\ y), \tag{5}$$

$$y''=f_x'(x,\ y)+f_y'(x,\ y)\,y'. \tag{6}$$

Процесс численного интегрирования дифференциального уравнения (1) при наличии начального условия (2), использующий

формулы (3) и (4), происходит следующим образом. Каким-либо методом вычисляем три начальные строки (*начальная таблица*):

$$x_0,\ y_0,\ y_0',\ y_0'';$$
$$x_1,\ y_1,\ y_1',\ y_1'';$$
$$x_2,\ y_2,\ y_2',\ y_2''.$$

Из формулы (4) при $i=2$ получаем первое приближение для y_3:

$$\bar{y}_3 = y_0 + 3(y_2 - y_1) + h^2(y_2'' - y_1'') \tag{7}$$

и, пользуясь формулами (5) и (6), находим для соответствующих производных y_3' и y_3'' их первые приближения:

$$\bar{y}_3' = f(x_3,\ \bar{y}_3) \text{ и } \bar{y}_3'' = f_x'(x_3,\ \bar{y}_3) + \bar{y}_3' f_y'(x_3,\ \bar{y}_3).$$

Второе приближение для y_3 определяем при $i=2$ из формулы (3):

$$\bar{\bar{y}}_3 = y_2 + \frac{h}{2}(y_2' + \bar{y}_3') + \frac{h^2}{12}(y_2'' - \bar{y}_3''). \tag{8}$$

После этого исправляем значения производных y_3' и y_3'', подсчитывая их вторые приближения:

$$\bar{\bar{y}}_3' = f(x_3,\ \bar{\bar{y}}_3) \text{ и } \bar{\bar{y}}_3'' = f_x'(x_3,\ \bar{\bar{y}}_3) + \bar{\bar{y}}_3' f_y'(x_3,\ \bar{\bar{y}}_3).$$

Для контроля еще раз вычисляем по формуле (3) третье приближение $\bar{\bar{\bar{y}}}_3$ значения y_3, используя найденные значения $\bar{\bar{\bar{y}}}_3'$ и $\bar{\bar{\bar{y}}}_3''$.

Если шаг h выбран подходящим, то пересчет не дает нового результата, и в этом случае можно положить $y_3 = \bar{\bar{\bar{y}}}_3$, $y_3' = \bar{\bar{\bar{y}}}_3'$, $y_3'' = \bar{\bar{\bar{y}}}_3''$.

В противном случае следует уменьшить шаг. Аналогично находятся дальнейшие значения y_i, y_i', y_i'' при $i>3$.

Для получения начальных значений y_1 и y_2 обычно используют метод последовательных приближений или метод Рунге — Кутта, после чего нужные производные y_i' и y_i'' ($i=0,\ 1,\ 2$) определяются по формулам (5) и (6).

Можно также применить следующий прием: сначала, используя данное начальное значение $y(x_0) = y_0$, непосредственно вычисляем

$$y_0' = f(x_0,\ y_0) \text{ и } y_0'' = f_x'(x_0,\ y_0) + y_0' f_y'(x_0,\ y_0).$$

Тем самым будет заполнена 1-я строка начальной таблицы.

Далее на основании формулы Тейлора приближенно получаем

$$y_1 = y_0 + hy_0' + \frac{1}{2}h^2 y_0''$$

и, следовательно, можно будет найти

$$y_1' = f(x_1,\ y_1) \text{ и } y_1'' = f_x'(x_1,\ y_1) + y_1' f_y'(x_1,\ y_1).$$

Пользуясь этими данными, уточняем значение y_1 по формуле (3):

$$y_1 = y_0 + \frac{h}{2}(y_0' + y_1') + \frac{h^2}{12}(y_0'' - y_1'')$$

и затем перевычисляем значения y_1' и y_1''. Тем самым заполняем 2-ю строку начальной таблицы. Аналогично, исходя из 2-й строки, находим элементы y_2, y_2' и y_2'' последней, 3-й строки начальной таблицы.

Отметим, что если пересчеты элементов строк дают значительные расхождения, то этот прием не является надежным. В таком случае следует или уменьшить шаг h вычислений, или же обратиться к более точным методам.

Пример 1. Для дифференциального уравнения

$$y' = x^2 - y^2$$

найти несколько значений интеграла $y(x)$, удовлетворяющего начальному условию $y(0) = 1$.

Решение. Примем шаг равным $h = 0{,}1$. Из начального условия имеем $y_0 = 1$. Методом Рунге—Кутта находим

$$y_1 = y(0{,}1) = 0{,}909410,\quad y_2 = y(0{,}2) = 0{,}835786.$$

Отсюда, пользуясь формулами

$$y_i' = x_i^2 - y_i^2 \tag{9}$$

и

$$y_i'' = 2(x_i - y_i y_i')\quad (i = 0,\ 1,\ 2,\ 3,\ \ldots), \tag{10}$$

вычисляем:

$$\begin{aligned} y_0' &= -1{,}000000, & y_0'' &= 2{,}0000;\\ y_1' &= -0{,}817026, & y_1'' &= 1{,}6860;\\ y_2' &= -0{,}658538, & y_2'' &= 1{,}5008. \end{aligned}$$

Дальнейшие значения y_i $(i \geqslant 3)$ находим, используя описанный выше регулярный процесс. Например, для $i = 3$ из формулы (7) получаем

$$\bar{y}_3 = y_0 + 3(y_2 - y_1) + 0{,}1^2(y_2'' - y_1'') =$$
$$= 1{,}000 + 3\cdot(0{,}835786 - 0{,}909410) + 0{,}01\cdot(1{,}5008 - 1{,}6860) = 0{,}777276.$$

Подставив это значение в формулы (9) и (10), где положено $i = 3$, будем иметь

$$\bar{y}_3' = 0{,}3^2 - 0{,}777276^2 = -0{,}514158,$$
$$\bar{y}_3'' = 2\,[0{,}3 - 0{,}777276\cdot(-0{,}514158)] = 1{,}3994.$$

Значение y_3 уточняем, пользуясь формулой (8):

$$\bar{\bar{y}}_3 = y_2 + \frac{0{,}1}{2}\cdot(\bar{y}_3' + y_2') - \frac{0{,}01}{12}\cdot(\bar{y}_3'' - y_2'') =$$

$$= 0{,}835786 + 0{,}05\cdot(-0{,}514158 - 0{,}658538) - \frac{0{,}01}{12}\cdot(1{,}3994 - 1{,}5008) = 0{,}777236.$$

Из формул (9) и (10) находим соответственно

$$\bar{\bar{y}}_3' = -0{,}514096 \quad \text{и} \quad \bar{\bar{y}}_3'' = 1{,}3991.$$

Полученное приближение $\bar{\bar{y}}_3$ подвергаем для уточнения итерационной обработке, подставляя его снова в формулу (8):

$$\bar{\bar{\bar{y}}}_3 = y_2 + \frac{0{,}1}{2}(\bar{\bar{y}}_3' + y_2') - \frac{0{,}01}{12}(\bar{\bar{y}}_3'' - y_2'') =$$

$$= 0{,}835786 + 0{,}05\cdot(-0{,}514096 - 0{,}658538) - \frac{0{,}01}{12}\cdot(1{,}3991 - 1{,}5008) = 0{,}777239.$$

Пользуясь формулами (9) и (10), получим соответственно

$$\bar{\bar{\bar{y}}}_3' = 0{,}514100 \quad \text{и} \quad \bar{\bar{\bar{y}}}_3'' = 1{,}3992.$$

Так как расхождение второго и третьего приближений незначительно, то принимаем:

$$y_3 = 0{,}777239, \qquad y_3' = 0{,}514100, \qquad y_3'' = 1{,}3992,$$

причем последние десятичные знаки приведенных значений сомнительны.

Аналогично вычисляются дальнейшие значения y_4, y_5, ...

Результаты окончательных вычислений помещены в таблице 57.

Таблица 57

Интегрирование дифференциального уравнения методом производных высших порядков

x	y	y'	y''
0	1,000000	−1,000000	2,00000
0,1	0,909410	−0,817026	1,6860
0,2	0,835786	−0,658538	1,5008
0,3	0,777239	−0,514100	1,3992
0,4	0,732728	−0,376890	1,3523
0,5	0,701769	−0,242480	1,3403
0,6	0,684230	−0,108169	1,3480

В заключение приведем формулы, обеспечивающие более высокую степень точности, но требующие вычисления, кроме второй, еще и третьей производной искомого решения. А именно, используя формулу Тейлора и употребляя прием, аналогичный указанному выше, получаем формулы

$$y_{i+1} = y_{i-2} + 3(y_i - y_{i-1}) + \frac{h^3}{2}(y_i''' + y_{i-1}''') + R_1, \qquad (11)$$

где $R_1 = \frac{420h^7 y^{(7)}(\xi_1)}{100800}$, и

$$y_{i+1} = y_i + \frac{h}{2}(y'_{i+1} + y'_i) - \frac{h^2}{10}(y''_{i+1} - y''_i) + \frac{h^3}{120}(y'''_{i+1} + y'''_i) + R_2, \quad (12)$$

где $R_2 = -\frac{h^7 y^{(7)}(\xi_2)}{100800}$.

Формула (11) употребляется для нахождения первого приближения $\bar{y}_{i+1}$; формула (12) дает уточненное значение $\bar{\bar{y}}_{i+1}$. Само собою разумеется, что к последним двум формулам целесообразно прибегать тогда, когда форма дифференциального уравнения позволяет сравнительно просто находить вторую и третью производные от искомой функции y.

§ 12. Численное интегрирование дифференциальных уравнений второго порядка

Задача Коши для дифференциального уравнения n-го порядка

$$y^{(n)} = f(x, y, y', \ldots, y^{(n-1)}) \quad (1)$$

при начальных условиях $y^{(k)}(x_0) = y_0^{(k)}$ $(k = 0, 1, 2, \ldots, n-1)$, как было показано в § 1, сводится к задаче Коши для системы

$$\left.\begin{aligned} y' &= y_1, \\ y'_1 &= y_2, \\ &\ldots\ldots\ldots\ldots \\ y'_{n-1} &= f(x, y, y_1, \ldots, y_{n-1}), \end{aligned}\right\} \quad (2)$$

где $y_k(x_0) = y_0^{(k)}$ $(k = 0, 1, 2, \ldots, n-1;\ y_0 \equiv y)$.

Поэтому изложенные методы приближенного интегрирования систем дифференциальных уравнений применимы также к уравнению (1).

Однако общие схемы для приближенного решения дифференциальных систем, не учитывающие специфических особенностей системы (2), оказываются излишне сложными. Поэтому целесообразно вывести формулы, специально приспособленные для численного интегрирования дифференциального уравнения вида (1). Мы ограничимся рассмотрением дифференциального уравнения второго порядка

$$y'' = f(x, y, y') \quad (3)$$

при начальных условиях $y(x_0) = y_0$, $y'(x_0) = y'_0$.

Выведем формулы для приближенного вычисления интеграла $y = y(x)$ дифференциального уравнения (3) с помощью метода Адамса.

Для этого выберем шаг $\Delta x = h$ и введем стандартные обозначения:

$$x_i = x_0 + ih, \quad y_i = y(x_i), \quad y'_i = y'(x_i),$$
$$y''_i = y''(x_i) = f(x_i, y_i, y'_i) \quad (i = 0, 1, 2, \ldots).$$

Допустим, что известны значения

$$y_0,\ y_1,\ \ldots,\ y_i;$$
$$y'_0,\ y'_1,\ \ldots,\ y'_i;$$
$$y''_0,\ y''_1,\ \ldots,\ y''_i$$

($i \geqslant 3$). Тогда можно вычислить разности $\Delta y''_{i-1}$, $\Delta^2 y''_{i-2}$, $\Delta^3 y''_{i-3}$.

Применяя вторую интерполяционную формулу Ньютона, с точностью до разностей четвертого порядка будем иметь (см. § 8)

$$y'' = y''_i + q\Delta y''_{i-1} + \frac{q(q+1)}{2!}\Delta^2 y''_{i-2} + \frac{q(q+1)(q+2)}{3!}\Delta^3 y''_{i-3}, \tag{4}$$

где

$$q = \frac{x - x_i}{h}. \tag{5}$$

Так как в силу формулы (5) $dx = h\,dq$, то, очевидно, имеем

$$y' = y'_i + \int_{x_i}^{x} y''\,dx = y'_i + h\int_0^q y''\,dq \tag{6}$$

и

$$y = y_i + \int_{x_i}^{x} y'\,dx = y_i + h\int_0^q y'\,dq. \tag{7}$$

Интегрируя последовательно на отрезке $[0, q]$ два раза по q формулу (4), на основании формул (6) и (7) получим

$$y' - y'_i = h\left[qy''_i + \frac{q^2}{2}\Delta y''_{i-1} + \left(\frac{q^2}{4} + \frac{q^3}{6}\right)\Delta^2 y''_{i-2} + \right.$$
$$\left. + \left(\frac{q^2}{6} + \frac{q^3}{6} + \frac{q^4}{24}\right)\Delta^3 y''_{i-3}\right] \tag{6'}$$

и

$$y - y_i = hqy'_i + h^2\left[\frac{q^2}{2}y''_i + \frac{q^3}{6}\Delta y''_{i-1} + \right.$$
$$\left. + \left(\frac{q^3}{12} + \frac{q^4}{24}\right)\Delta^2 y''_{i-2} + \left(\frac{q^3}{18} + \frac{q^4}{24} + \frac{q^5}{120}\right)\Delta^3 y''_{i-3}\right]. \tag{7'}$$

Отсюда, полагая $q = 1$ в формулах (6′) и (7′), находим

$$\Delta y'_i = h\left(y''_i + \frac{1}{2}\Delta y''_{i-1} + \frac{5}{12}\Delta^2 y''_{i-2} + \frac{3}{8}\Delta^3 y''_{i-3}\right) \tag{8}$$

и

$$\Delta y^1_i = hy'_i + h^2\left(\frac{1}{2}y''_i + \frac{1}{6}\Delta y''_{i-1} + \frac{1}{8}\Delta^2 y''_{i-2} + \frac{1}{10}\Delta^3 y''_{i-3}\right), \tag{9}$$

где положено $\frac{19}{180} \approx \frac{1}{10}$. Можно принять

$$y'_{i+1} = y'_i + \Delta y'_i,\quad y_{i+1} = y_i + \Delta y^1_i. \tag{10}$$

Более точный вариант счета следующий [14]: найдя y'_{i+1}, по формулам (8) и (10) вычисляют $\Delta y'_i$, $\Delta^2 y'_{i-1}$, $\Delta^3 y'_{i-2}$, после чего определяют Δy_i по более точной формуле (см. § 8):

$$\Delta y_i^{\text{II}} = h\left(y'_i + \frac{1}{2}\Delta y'_i - \frac{1}{12}\Delta^2 y'_{i-1} - \frac{1}{24}\Delta^3 y'_{i-2}\right).$$

Затем, приняв

$$y_{i+1} = y_i + \Delta y_i^{\text{II}}, \quad y'_{i+1} = y'_i + \Delta y_i^{\text{I}}, \tag{11}$$

из дифференциального уравнения можно найти

$$y''_{i+1} = f(x_{i+1}, y_{i+1}, y'_{i+1})$$

и, пополнив таблицу, вычислить y'_{i+1} по формуле

$$y'_{i+1} = y'_i + \Delta y_i^{\text{II}},$$

где

$$\Delta y_i^{\text{II}} = h\left(y''_i + \frac{1}{2}\Delta y''_i - \frac{1}{12}\Delta^2 y''_{i-1} - \frac{1}{24}\Delta^3 y''_{i-2}\right). \tag{12}$$

В случае необходимости повторяют аналогичный пересчет величин y_{i+1} и y'_{i+1} до тех пор, пока не прекратятся изменения. Рекомендуется шаг h выбирать столь малым, чтобы формулы (10) и (11) давали одинаковые результаты в пределах заданной точности.

Что касается начального отрезка y_0, y_1, y_2, y_3; y'_0, y'_1, y'_2, y'_3, то он предварительно определяется каким-нибудь подходящим методом.

Разработаны также другие приемы для приближенного интегрирования дифференциального уравнения (3) (см. [14], [18], [27]). В частности, для дифференциального уравнения вида

$$y'' = f(x, y)$$

имеется весьма точный метод Б. В. Нумерова [27].

Пример 1. На отрезке [0, 1] найти интеграл $y = y(x)$ уравнения

$$y'' + y \operatorname{ch} x = 0, \tag{13}$$

удовлетворяющий начальным условиям

$$y(0) = 0, \quad y'(0) = 1. \tag{14}$$

Решение. Примем шаг $h = 0{,}2$. Для подсчета начального отрезка применим метод степенных рядов. Имеем

$$y'' = -y \operatorname{ch} x;$$

отсюда

$$y''' = -(y \operatorname{sh} x + y' \operatorname{ch} x),$$
$$y^{\text{IV}} = -(y \operatorname{ch} x + 2y' \operatorname{sh} x + y'' \operatorname{ch} x),$$
$$y^{\text{V}} = -(y \operatorname{sh} x + 3y' \operatorname{ch} x + 3y'' \operatorname{sh} x + y''' \operatorname{ch} x) \text{ и т. д.}$$

Следовательно, в силу начальных условий (14) получаем:

$$y_0=0,\quad y_0'=1,\quad y_0''=0,\quad y_0'''=-1,$$
$$y_0^{IV}=0,\quad y_0^{V}=-2,\ \ldots,$$

поэтому

$$y=x-\frac{x^3}{6}-\frac{x^5}{60}+\ldots \text{ и } y'=1-\frac{x^2}{2}-\frac{x^4}{12}+\ldots \tag{15}$$

Полагая $x_i=0{,}2i$ $(i=-1,\ 0,\ 1,\ 2)$, из формул (15) с точностью до 10^{-3} находим:

$$y_1=-y_{-1}=0{,}200-0{,}001=0{,}199,$$
$$y_1'=y_{-1}'=1-0{,}020=0{,}980,$$
$$y_2=0{,}400-0{,}011=0{,}389,$$
$$y_2'=1-0{,}080-0{,}002=0{,}918.$$

Дальнейшие вычисления производим по формулам (8), (9), (10) без пересчета. Результаты вычислений приведены в таблицах 58 и 59.

Таблица 58

Основной бланк для решения задачи Коши (13) — (14)

i	x	y	Δy	y'	$\Delta y'$	ch x	y''	$\Delta y''$	$\Delta^2 y''$	$\Delta^3 y''$
−1	−0,2	−0,199		0,980		1,020	0,203	−203	0	−15
0	0	0		1		1	0	−203	−15	−13
1	0,2	0,199		0,980		1,020	−0,203	−218	−28	− 8
2	0,4	0,389	0,174	0,918	−0,108	1,081	−0,421	−246	−36	
3	0,6	0,563	0,147	0,810	−0,161	1,185	−0,667	−282		
4	0,8	0,710	0,109	0,649		1,337	−0,949			
5	1,0	0,819								

Таблица 59

Вспомогательный бланк для решения задачи Коши (13) — (14)

i	2	3	4
$y_i^{(1)}$	−0,421	−0,667	
$\frac{1}{2}\Delta y_{i-1}''$	−0,109	−0,123	
$\frac{5}{12}\Delta^2 y_{i-2}''$	−0,006	−0,012	
$\frac{3}{8}\Delta^3 y_{i-3}''$	−0,006	−0,005	
Σ_1	−0,542	−0,807	
$\Delta y_i'=h\Sigma_1$	−0,108	−0,161	

Продолжение табл. 59

i	2	3	4
$\frac{1}{2}y_i''$	−0,210	−0,334	−0,474
$\frac{1}{6}\Delta y_{i-1}''$	36	−0,041	−0,047
$\frac{1}{8}\Delta^2 y_{i-2}''$	2	−0,004	−0,004
$\frac{1}{10}\Delta^3 y_{i-3}''$	2	−0,001	−0,001
Σ_2	−0,250	−0,380	−0,526
$h^2\Sigma_2$	−0,010	−0,015	−0,021
hy_i'	+0,184	+0,162	+0,130
Δy_i	0,174	0,147	0,109

§ 13. Метод Чаплыгина

Рассмотренные выше приближенные методы интегрирования дифференциальных уравнений были преимущественно численными. В этом параграфе мы изложим аналитический метод приближенного интегрирования дифференциальных уравнений, принадлежащий Чаплыгину и основанный на совершенно новой идее.

Метод Чаплыгина [15], [16], [5], [2] является одним из наиболее точных аналитических методов приближенного интегрирования дифференциальных уравнений и в то же время допускающим простую оценку погрешности решения. Сущность его состоит в том, что искомое решение $y=y(x)$ $(x_0 \leqslant x \leqslant X)$ аппроксимируется двумя последовательностями функций

$$u_n = u_n(x) \text{ и } v_n = v_n(x) \quad (n=1,\ 2,\ 3,\ \ldots),$$

удовлетворяющими двойному неравенству

$$u_n(x) \leqslant y(x) \leqslant v_n(x) \text{ при } x \in [x_0,\ X]$$

и начальному условию

$$u_n(x_0) = v_n(x_0) = y(x_0),$$

причем такими, что $v_n(x) - u_n(x) \rightrightarrows 0$ на $[x_0,\ X]$ при $n \to \infty$. Геометрически это значит, что искомая интегральная кривая $y=y(x)$

зажимается в сколь угодно узкий криволинейный сектор $A_0B_nC_n$ (рис. 33).

Таким образом, для решения $y(x)$ строится «вилка» $[u_n(x), v_n(x)]$ и указывается регулярный процесс, с помощью которого можно сузить эту «вилку» до желаемых размеров. В этом смысле метод Чаплыгина напоминает комбинированный метод решения обычных уравнений (см. [23], гл. IV).

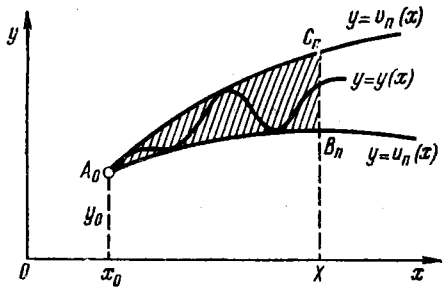

Рис. 33.

Если положить $y_n \approx u_n(x)$, то предельная абсолютная погрешность приближенного решения $u_n(x)$ будет равна

$$\varepsilon_n(x) = v_n(x) - u_n(x),$$

т. е. эта погрешность на каждом шаге определяется непосредственно.

Для простоты изложим идею метода Чаплыгина применительно к дифференциальному уравнению первого порядка

$$y' = f(x, y) \tag{1}$$

с начальным условием

$$y(x_0) = y_0, \tag{2}$$

причем будем предполагать, что правая часть $f(x, y)$ непрерывна и имеет непрерывные производные $f'_y(x, y)$ и $f''_{yy}(x, y)$ в некоторой окрестности начальной точки $M_0(x_0, y_0)$. Метод основан на одной лемме, представляющей также самостоятельный математический интерес.

Лемма Чаплыгина об интегральных неравенствах. *Пусть*

$$I[z] = z' - f(x, z)$$

— дифференциальный оператор, соответствующий дифференциальному уравнению (1), *и* $y = y(x)$ *— интеграл уравнения* (1), *т. е.*

$$I[y] = 0, \tag{3}$$

удовлетворяющий начальному условию $y(x_0) = y_0$ *и определенный при* $x_0 \leqslant x \leqslant X$.

Если функция $u = u(x) \in C^1$ *удовлетворяет условиям:*

$$I[u] \leqslant 0 \text{ при } x_0 \leqslant x \leqslant X \tag{4}$$

и

$$u(x_0) = y_0,$$

то на отрезке $[x_0, X]$ *выполнено неравенство*

$$u \leqslant y, \tag{5}$$

т. е. функция u *является нижним приближением решения* y.

Аналогично, если для функции $v = v(x) \in C^1$ *выполнены условия:*

$$I[v] \geqslant 0 \text{ при } x_0 \leqslant x \leqslant X \tag{6}$$

и

$$v(x_0) = y_0,$$

то на отрезке $[x_0, X]$ *имеет место неравенство*

$$y \leqslant v, \tag{7}$$

т. е. функция v *является верхним приближением решения* y.

Доказательство. Очевидно, достаточно доказать лишь одно из неравенств (5) или (7). Докажем, например, неравенство (5).

Из формул (3) и (4) имеем $y' - f(x, y) = 0$ и $u' - f(x, u) \leqslant 0$. Отсюда

$$(y - u)' - p(x)(y - u) \geqslant 0, \tag{8}$$

где

$$p(x) = \frac{f(x, y(x)) - f(x, u(x))}{y - u}. \tag{9}$$

Функция $p(x)$ теряет смысл при x, для которого $y = u$. В этом случае полагают

$$p(x) = \lim_{u \to y} \frac{f(x, y) - f(x, u)}{y - u} = f'_y(x, y(x)).$$

В силу приведенных выше условий функция $p(x)$ определена и непрерывна на отрезке $[x_0, X]$.

Умножим обе части дифференциального неравенства (8) на положительный интегрирующий множитель

$$\mu(x) = e^{\int_{x_0}^{x} p(x)\,dx};$$

будем иметь

$$\frac{d}{dx}\{\mu(x)[y(x) - u(x)]\} \geqslant 0. \tag{10}$$

Отсюда, интегрируя неравенство (10) в пределах от x_0 до x, где $x_0 \leqslant x \leqslant X$, получим

$$\mu(x)[y(x) - u(x)] - \mu(x_0)[y(x_0) - u(x_0)] \geqslant 0$$

или, так как $y(x_0) = u(x_0)$ и $\mu(x) > 0$, то окончательно находим

$$u(x) \leqslant y(x) \text{ при } x_0 \leqslant x \leqslant X,$$

что и требовалось доказать.

Следствие. Пусть функции $y=y(x)$, $u=u(x)$ и $v=v(x)$ при $x_0 \leqslant x \leqslant X$ удовлетворяют соответственно дифференциальным уравнениям

$$y'=f(x, y), \tag{11}$$

$$u'=f_1(x, u), \quad v'=f_2(x, v), \tag{12}$$

где

$$f_1(x, z) \leqslant f(x, z) \leqslant f_2(x, z), \tag{13}$$

и общему начальному условию $y(x_0)=u(x_0)=v(x_0)=y_0$. Тогда справедливо неравенство

$$u(x) \leqslant y(x) \leqslant v(x) \quad \text{при} \quad x_0 \leqslant x \leqslant X. \tag{14}$$

Действительно, полагая $I[z]=z'-f(x, z)$, из условий (12) и (13) имеем

$$I[u]=u'-f(x, u)=f_1(x, u)-f(x, u) \leqslant 0,$$
$$I[v]=v'-f(x, v)=f_2(x, v)-f(x, v) \geqslant 0.$$

Отсюда на основании предыдущей леммы следует неравенство (14).

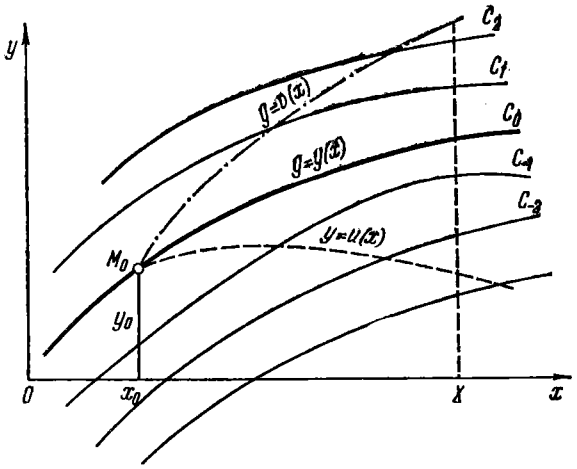

Рис. 34.

Нетрудно выяснить геометрический смысл неравенства (13).

Рассмотрим семейство интегральных кривых $\ldots, C_{-2}, C_{-1}, C_0, C_1, C_2, \ldots$ (K) уравнения (11), где C_0 — интегральная кривая, проходящая через точку $M_0(x_0, y_0)$, т. е. график решения $y=y(x)$ (рис. 34).

Из первой части неравенства (13) вытекает, что интегральная кривая первого из уравнений (12) $y=u(x)$, проходящая через точку M_0, пересекая последовательные кривые C семейства (K), опу-

скается по отношению к этим кривым и, следовательно, лежит ниже кривой C_0 (рис. 34).

Наоборот, из второй части неравенства (13) усматриваем, что интегральная кривая второго из уравнений (12) $y=v(x)$, проходящая через точку M_0, пересекая последовательно кривые C семейства (K), поднимается по отношению к этим кривым и поэтому расположена выше кривой C_0 (рис. 34).

Таким образом, имеет место неравенство (14).

Покажем теперь метод построения последовательностей функций u_n, v_n $(n=0, 1, 2, \ldots)$, аппроксимирующих решение y дифференциального уравнения (1).

Предположим сначала, что $\frac{\partial^2 f}{\partial y^2}$ сохраняет постоянный знак в рассматриваемой области, например:

$$\frac{\partial^2 f}{\partial y^2} \geqslant 0. \tag{15}$$

Методом проб подберем две функции $u_0=u_0(x)$ и $v_0=v_0(x)$, где $u_0(x) \leqslant v_0(x)$, для которых выполняются соответственно неравенства

$$I[u_0]=-\varphi_0(x) \leqslant 0 \quad \text{и} \quad I[v_0]=\psi_0(x) \geqslant 0,$$

где $\varphi_0(x)$ и $\psi_0(x)$ — соответствующие *невязки*.

В случае затруднений при нахождении функций u_0 и v_0 можно применить следующий прием. Пусть правая часть $f(x, y)$ дифференциального уравнения (1) определена и непрерывна в некоторой прямоугольной области $R\{x_0 \leqslant x \leqslant x_0+a,\ |y-y_0| \leqslant b\}$, где a и b — положительные числа. Положим

$$m=\min f(x, y) \text{ при } (x, y) \in R;\ M=\max f(x, y) \text{ при } (x, y) \in R; \tag{16}$$

$$h=\min\left\{a,\ \frac{b}{|m|},\ \frac{b}{|M|}\right\}. \tag{17}$$

Тогда на отрезке $[x, X]$, где $X=x_0+h$, в качестве начальных приближений можно выбрать следующие функции (рис. 35):

$$u_0=y_0+m(x-x_0), \quad v_0=y_0+M(x-x_0).$$

Действительно, из формул (16) и (17) вытекает, что

$$I[u_0]=m-f(x, u_0) \leqslant 0, \quad I[v_0]=M-f(x, v_0) \geqslant 0$$

при $x_0 \leqslant x \leqslant x_0+h$.

Для уточнения приближений u_0 и v_0 полагаем $u_1=u_0+\rho_0$ и $v_1=v_0-\sigma_0$, где $\rho_0=\rho_0(x)$ и $\sigma_0=\sigma_0(x)$ — функции, определяемые из линейных дифференциальных уравнений

$$\frac{d\rho_0}{dx}=p_0(x)\rho_0+\varphi_0(x), \quad \rho_0(x_0)=0 \tag{18'}$$

и

$$\frac{d\sigma_0}{dx} = q_0(x)\,\sigma_0 + \psi_0(x), \quad \sigma_0(x_0) = 0, \tag{18''}$$

где $\varphi_0(x) = -I[u_0]$, $\psi_0(x) = I[v_0]$,

$$p_0(x) = f'_y(x, u_0), \quad q_0(x) = \frac{f(x, v_0) - f(x, u_0)}{v_0 - u_0}. \tag{19}$$

Так как свободные члены $\varphi_0(x)$ и $\psi_0(x)$ линейных уравнений (18′)

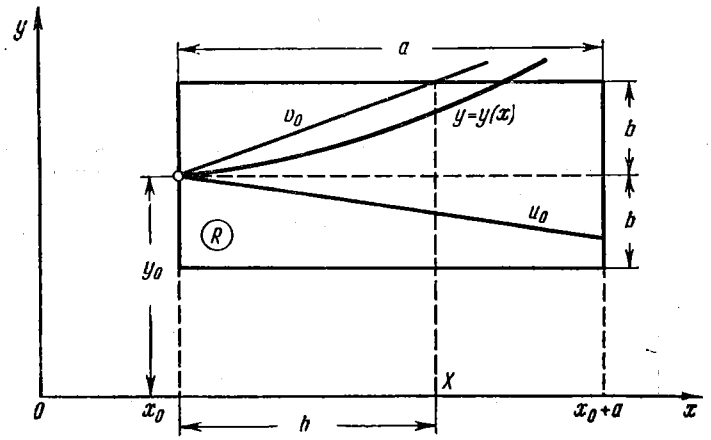

Рис. 35.

и (18″) неотрицательны, то из известных общих формул для решений этих уравнений следует, что

$$\rho_0(x) \geqslant 0 \text{ и } \sigma_0(x) \geqslant 0 \text{ при } x_0 \leqslant x \leqslant X.$$

Можно доказать [16], что

$$I[u_1] \leqslant 0 \quad \text{и} \quad I[v_1] \geqslant 0 \quad \text{на} \quad [x_0, X].$$

Таким образом, u_1 и v_1 являются соответственно нижним и верхним приближениями для точного решения y, образующими более узкую «вилку», чем начальные функции u_0 и v_0.

Вообще, если $I[u_n] = -\varphi_n(x) \leqslant 0$ и $I[v_n] = \psi_n(x) \geqslant 0$, то можно положить

$$u_{n+1} = u_n + \rho_n \quad \text{и} \quad v_{n+1} = v_n - \sigma_n,$$

причем функции ρ_n и σ_n неотрицательны и определяются из линейных дифференциальных уравнений

$$\frac{d\rho_n}{dx} = p_n(x)\,\rho_n + \varphi_n(x), \quad \rho_n(x_0) = 0 \tag{20'}$$

и

$$\frac{d\sigma_n}{dx} = q_n(x)\,\sigma_n + \psi_n(x_0), \quad \sigma_n(x_0) = 0 \quad (n = 1, 2, 3, \ldots), \tag{20''}$$

где

$$p_n(x) = f'_y(x, u_n), \quad q_n(x) = \frac{f(x, v_n) - f(x, u_n)}{v_n - u_n}.$$

Доказывается, что

$$I[u_{n+1}] \leqslant 0 \quad \text{и} \quad I[v_{n+1}] \geqslant 0. \tag{21}$$

Отсюда получаем (рис. 36)

$$u_0 \leqslant u_1 \leqslant u_2 \leqslant \ldots \leqslant u_n \leqslant u_{n+1} \leqslant \ldots \leqslant y$$

и

$$y \leqslant \ldots \leqslant v_{n+1} \leqslant v_n \leqslant \ldots \leqslant v_1 \leqslant v_0.$$

Заметим, что если зафиксировать x и построить график функции

$$z = f(x, y), \tag{22}$$

то для каждого значения x получим, что $p_n(x) = \operatorname{tg} \alpha_n$ представляет

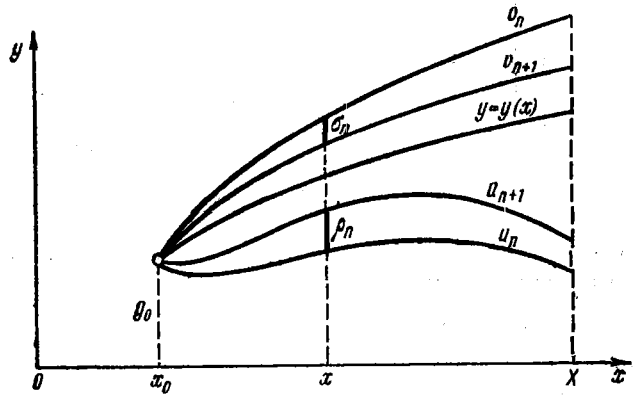

Рис. 36.

собой угловой коэффициент касательной к кривой (22) в точке $y = u_n$, а $q_n(x) = \operatorname{tg} \beta_n$ есть угловой коэффициент секущей к кривой (22), проходящей через точки $y = u_n$ и $y = v_n$ (рис. 37).

Поэтому получение u_n условно может быть названо методом касательных, а получение v_n — методом хорд.

Из уравнений (20′) и (20″) будем иметь

$$\rho_n = e^{\int_{x_0}^{x} p_n(x)\,dx} \int_{x_0}^{x} \varphi_n(x)\, e^{-\int_{x_0}^{x} p_n(x)\,dx}\, dx \tag{23}$$

и

$$\sigma_n = e^{\int_{x_0}^{x} q_n(x)\,dx} \int_{x_0}^{x} \psi_n(x)\, e^{-\int_{x_0}^{x} q_n(x)\,dx}\, dx \qquad (n=0,\ 1,\ 2,\ \ldots). \quad (24)$$

Н. Н. Лузин [16] доказал, что при $n \to \infty$ разность $v_n - u_n$ весьма быстро стремится к нулю. Если для некоторого n имеем $v_n - u_n < \varepsilon$, где ε — заданная предельная погрешность, то приближенно можно положить

$$y \approx \frac{1}{2}(u_n + v_n).$$

Заметим, что если не требуется оценки погрешности приближенного решения, то нет необходимости строить две последовательности

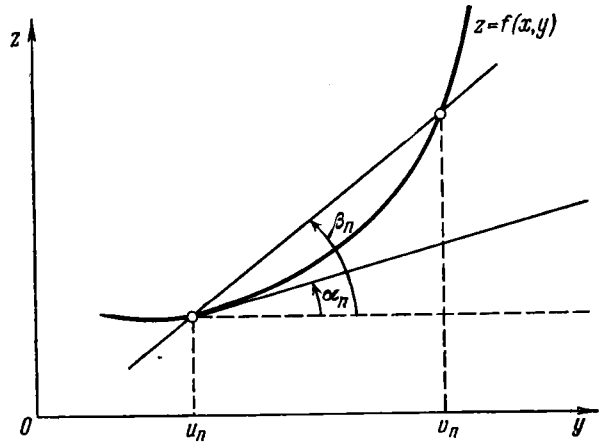

Рис. 37.

u_n и v_n $(n=1, 2, \ldots, n)$. В этом случае достаточно ограничиться последовательностью u_n (или v_n), так как

$$y = \lim_{n\to\infty} u_n.$$

Если вместо неравенства (15) выполнено обратное неравенство

$$\frac{\partial^2 f}{\partial y^2} \leqslant 0,$$

то функции u_n и v_n меняются своими местами.

Случай, когда $\frac{\partial^2 f}{\partial y^2}$ на отрезке $[x_0, X]$ меняет знак, более сложен, поэтому здесь его рассматривать не будем [15], [16].

Пример 1. Методом Чаплыгина построить на отрезке $\left[0, \frac{1}{2}\right]$ несколько приближений для интеграла *уравнения Рикатти*

$$y' = x + 2y^2, \quad y(0) = 0.$$

Решение. В нашем случае

$$f(x, y) = x + 2y^2 \quad \text{и} \quad I[z] = z' - x - 2z^2.$$

Так как $f''_{yy}(x, y) = 4 > 0$, то можно применить изложенную выше теорию. Для пробы положим

$$u_0 = 0 \quad \text{и} \quad v_0 = x.$$

Имеем

$$I[u_0] = -x = -\varphi_0(x) \leqslant 0,$$

$$I[v_0] = 1 - x - 2x^2 = \psi_0(x) \geqslant 0$$

при $0 \leqslant x \leqslant \frac{1}{2}$. Следовательно, получим $0 \leqslant y \leqslant x$ при $0 \leqslant x \leqslant 1/2$ (рис. 38).

Согласно общей теории полагаем

$$p_0(x) = f_y(x, u_0) = 4u_0 = 0$$

и

$$q_0(x) = \frac{f(x, v_0) - f(x, u_0)}{v_0 - u_0} = \frac{(x + 2v_0^2) - (x + 2u_0^2)}{v_0 - u_0} = 2(v_0 + u_0) = 2x.$$

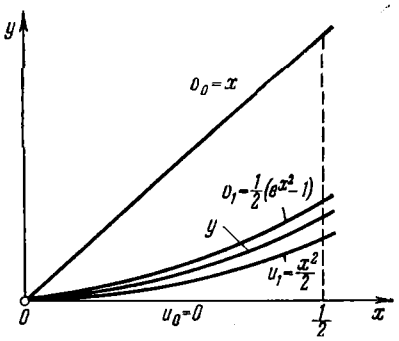

Рис. 38.

Пусть $u_1 = u_0 + \rho_0$ и $v_1 = v_0 - \sigma_0$. В силу формулы (23) находим

$$\rho_0 = e^{\int_0^x p_0(x)\,dx} \int_0^x \psi_0(x)\, e^{-\int_0^x p_0(x)\,dx}\, dx = \int_0^x x\,dx = \frac{x^2}{2}.$$

Поэтому $u_1 = \frac{x^2}{2}$. На основании формулы (24) получаем

$$\sigma_0 = e^{\int_0^x q_0(x)\,dx} \int_0^x \psi_0(x)\, e^{-\int_0^x q_0(x)\,dx}\, dx =$$

$$= e^{\int_0^x 2x\,dx} \int_0^x (1 - x - 2x^2)\, e^{-\int_0^x 2x\,dx}\, dx = e^{x^2} \int_0^x (1 - x - 2x^2)\, e^{-x^2} dx.$$

Интеграл

$$\Phi(x) = \int_0^x e^{-x^2}\, dx$$

является неэлементарной функцией, которая только числовым множителем отличается от интеграла вероятностей [17].

Очевидно,

$$\int_0^x (1-x-2x^2)\,e^{-x^2}\,dx=\int_0^x e^{-x^2}\,dx-\int_0^x xe^{-x^2}\,dx+\int_0^x x\,d\,(e^{-x^2})=$$

$$=\Phi(x)+\frac{1}{2}(e^{-x^2}-1)+xe^{-x^2}-\Phi(x)=-\frac{1}{2}(1-e^{-x^2})+xe^{-x^2}.$$

Поэтому

$$\sigma_0=-\frac{1}{2}e^{x^2}+\frac{1}{2}+x,$$

и, следовательно,

$$v_1=x-\left(\frac{1}{2}+x-\frac{1}{2}e^{x^2}\right)=\frac{1}{2}(e^{x^2}-1).$$

Для контроля произведем проверку знака оператора $l[z]$ при $z=u_1$ и $z=v_1$. Имеем

$$l[u_1]=x-x-2x^2=-2x^2\leqslant 0.$$

Далее,

$$l[v_1]=xe^{x^2}-x-\frac{1}{2}(e^{2x^2}-2e^{x^2}+1)=$$

$$=(1+x)(e^{x^2}-1)-\frac{1}{2}(e^{2x^2}-1)=$$

$$=(e^{x^2}-1)\left(1+x-\frac{e^{x^2}+1}{2}\right)=(e^{x^2}-1)\,\frac{1+2x-e^{x^2}}{2}.$$

При $0\leqslant x\leqslant 1/2$ на основании формулы Тейлора получаем

$$e^{x^2}=1+x^2e^{(\theta x)^2}<1+2x^2,$$

где $0<\theta<1$; поэтому

$$l[v_1]>(e^{x^2}-1)(x-x^2)\geqslant 0.$$

Оценим разность

$$v_1-u_1=\frac{1}{2}(e^{x^2}-1)-\frac{x^2}{2}.$$

Так как $(v_1-u_1)'=x(e^{x^2}-1)\geqslant 0$ при $0\leqslant x\leqslant 1/2$, то

$$\max(v_1-u_1)=\frac{1}{2}(e^{1/4}-1)-\frac{1}{8}=\frac{1}{2}(1{,}2840-1)-0{,}125=0{,}017.$$

Итак,

$$y\approx\frac{1}{4}(e^{x^2}-1+x^2)\ \text{с точностью до } 0{,}02.$$

Аналогично в случае необходимости можно получить и дальнейшие приближения искомого решения.

Замечание. Большим неудобством метода Чаплыгина является возможность появления неберущихся квадратур. В этом случае интегралы приходится вычислять приближенно.

§ 14. Метод Ньютона — Канторовича

Обобщая известный метод Ньютона для нахождения корней алгебраического или трансцендентного уравнения [23], Л. В. Канторович [30] указал способ решения задачи Коши, а также краевых задач для дифференциальных уравнений. Сущность этого способа заключается в нахождении решения дифференциального уравнения путем последовательных уточнений начального приближения, находимых из линейных дифференциальных уравнений.

Пусть, например, дано дифференциальное уравнение первого порядка

$$y' = f(x, y) \tag{1}$$

с начальным условием

$$y(x_0) = y_0, \tag{2}$$

где функция $f(x, y)$ дважды дифференцируема в области $R\{|x - x_0| \leqslant a, |y - y_0| \leqslant b\}$.

Предположим, что $y_0 = y_0(x)$ является приближенным решением уравнения (1), причем

$$y_0(x_0) = y_0.$$

Тогда дальнейшие приближения y_n $(n = 1, 2, \ldots)$ решения задачи Коши (1), (2) последовательно могут быть определены по формуле

$$y'_{n+1} - p_n(x) y_{n+1} = q_n(x), \tag{3}$$

где $p_n(x) = f'_y(x, y_n(x))$ и $q_n(x) = f(x, y_n(x)) - f'_y(x, y_n(x)) y_n(x)$. Отсюда

$$y_{n+1}(x) = e^{\int\limits_{x_0}^{x} p_n(t)\,dt} \left\{ y_0 + \int\limits_{x_0}^{x} q_n(\tau)\, e^{-\int\limits_{x_0}^{\tau} p_n(t)\,dt} d\tau \right\},$$

$$(n = 0, 1, 2, \ldots).$$

Доказывается [30], что если $y_n(x) \in R$ $(n = 1, 2, \ldots)$, то

$$y(x) = \lim_{n \to \infty} y_n(x)$$

при $|x - x_0| \leqslant h$, где h достаточно мало, причем быстрота сходимости оценивается неравенством

$$|y_n(x) - y(x)| \leqslant \frac{1}{2^{n-1}} \gamma^{2^n - 1} \rho,$$

где

$$\rho = \max_{|x - x_0| \leqslant h} \left| y_0(x) - y_0 - \int\limits_{x_0}^{x} f(x, y_0(t))\,dt \right|, \quad \gamma = 2h\rho M_2 e^{2M_1 h}$$

и

$$|f'_y(x, y)| \leqslant M_1, \quad |f''_{yy}(x, y)| \leqslant M_2 \text{ в } R.$$

Метод быстро сходится на промежутке $|x - x_0| \leqslant h$, где $\gamma < 1$.

§ 15. Некоторые замечания об оценке погрешностей решений дифференциальных уравнений

Основная цель приближенных вычислений заключается в нахождении нужного результата с заданной степенью точности. Для этого прежде всего необходимо, чтобы все произведенные выкладки не содержали принципиальных и арифметических ошибок. Последнее обстоятельство особенно важно подчеркнуть для численных методов решения дифференциальных уравнений, где схемы вычислений достаточно сложны и малейшая ошибка безнадежно портит всю дальнейшую работу.

Если вычислитель по косвенным признакам заподозрит наличие такой ошибки, то для поисков и устранения ее требуется много времени. Поэтому разумно проводить превентивную политику, т. е. не допускать появления таких ошибок и продолжать счет лишь в том случае, если имеется абсолютная уверенность в правильности всех предшествующих вычислений.

Практика показывает, что в счете следует избегать поспешности и придерживаться золотого правила: «лучше меньше, да лучше».

Опытный вычислитель знает, что нельзя верить своим однократно проведенным выкладкам. Необходим действенный текущий контроль над процессом вычислений, позволяющий немедленно обнаружить возникшие ошибки. Для этой цели служат различного рода контрольные формулы и вспомогательные соображения. Начинающий вычислитель часто считает контроль излишним и применение контрольных формул неприятной дополнительной нагрузкой. Однако на своем печальном опыте он скоро убеждается, что поиски допущенной ошибки требуют большей затраты времени, значительно превосходящей время, необходимое на контроль.

Гарантией правильности решения может служить:

1) проверка выполнения условий задачи (например, для дифференциального уравнения найденное приближенное решение можно подставить в это уравнение или эквивалентное ему и проверить расхождение правой и левой частей);

2) двойной пересчет по возможности другим методом или другим вычислителем;

3) применение более грубой схемы и качественный анализ задачи.

Если известен алгоритм нахождения точного решения задачи, то других ошибок, кроме вычислительных, не имеется. Положение услож-

няется, если для нахождения решения приходится пользоваться приближенным методом. Здесь наряду со случайными вычислительными ошибками появляются неизбежные погрешности приближенного решения вообще. Отметим, что оценка погрешности приближенного решения или общие соображения о его характере существенно необходимы, так как в противном случае полученный формально правильный результат может иметь чисто иллюзорное значение.

Для иллюстрации приведем пример [18]. Пусть требуется численно найти решение задачи Коши:

$$y'' = 10y' + 11y, \qquad y(0) = 1; \; y'(0) = -1 \tag{1}$$

на отрезке [0, 3].

Общее решение дифференциального уравнения (1) имеет вид

$$y = c_1 e^{-x} + c_2 e^{11x},$$

где c_1 и c_2 — произвольные постоянные.

Для искомого решения эти постоянные имеют значения $c_1 = 1$ и $c_2 = 0$, поэтому $y = e^{-x}$, и, следовательно,

$$y(3) = e^{-3} \approx 0{,}0498. \tag{2}$$

Применение обычных приближенных методов, в силу неизбежных погрешностей округления, даст приближенное решение вида

$$\tilde{y} = (1 + \varepsilon_1) e^{-x} + \varepsilon_2 e^{11x},$$

где ε_1 и ε_2 — малые по абсолютной величине постоянные.

Пусть, например, точность метода такова, что $\varepsilon_1 = \pm 10^{-10}$ и $\varepsilon_2 = \pm 10^{-10}$. Тогда

$$\tilde{y}(3) = (1 \pm 10^{-10}) e^{-3} \pm 10^{-10} e^{33} \approx \pm 2 \cdot 10^4, \tag{3}$$

т. е. даже при применении приближенного метода высокой точности полученное значение (3) может не иметь ничего общего с точным значением (2).

На практике примеры такого рода встречаются сравнительно редко, однако с возможностью их приходится считаться. В сомнительных случаях рекомендуется составлять общее представление о поведении интегральных кривых, в зависимости от чего и выбирать подходящий способ вычисления.

При дальнейших рассуждениях будем предполагать, что вычислительные ошибки тем или иным способом устранены полностью. Займемся сейчас анализом ошибок, источником которых является приближенность вычислений.

Общая схема применения приближенных методов такова. Пусть $y(x)$ — искомая функция, значения $y(x_i)$ которой обычно определяются для системы равноотстоящих точек $x_i = x_0 + ih$ $(i = 0, 1, 2, \ldots)$.

Обозначим через $\tilde{y}_i = \tilde{y}(x_i, \tilde{y}_0, \tilde{y}_1, \ldots, \tilde{y}_{i-1})$ последовательные значения приближенного решения. Тогда предельная абсолютная погрешность приближенного решения $\tilde{y}$ выразится следующим образом:

$$E_i \geqslant |\tilde{y}_i - y_i|. \tag{4}$$

Пусть $\tilde{y}_i^*$ есть результат применения приближенного метода к точному решению, т. е. $\tilde{y}_i^* = \tilde{y}(x_i, y_0, y_1, \ldots, y_{i-1})$. Тогда можно принять $E_i = E_i^* + \varepsilon_i$, где

$$E_i^* = |\tilde{y}_i^* - y_i| \tag{5}$$

— так называемая *погрешность метода*, а

$$\varepsilon_i = |\tilde{y}_i - \tilde{y}_i^*| \tag{6}$$

— так называемая *текущая погрешность*.

Погрешность метода E_i^* представляет собой ошибку, происходящую от замены точного алгоритма решения приближенным. Эта погрешность неустранима. Поэтому метод вычислений должен быть выбран так, чтобы погрешность его на последнем шаге вычислений не превышала заданной величины.

Что касается текущей погрешности метода ε_i, то источником ее является расхождение между результатами применения данного метода к точному и приближенному решениям. Значительную роль здесь играет округление промежуточных данных.

Обычно порядок точности приближенного метода бывает известен. Под этим понимается следующее. Пусть известны значения точного решения $y_0, y_1, y_2, \ldots, y_i$ и очередное значение y_{i+1} разложено по степеням шага h в ряд Тейлора, т. е.

$$y_{i+1} = a_0^{(i)} + a_1^{(i)} h + a_2^{(i)} h^2 + \ldots,$$

где $a_p^{(i)} = \frac{1}{p!} y_i^{(p)}$ $(p = 0, 1, 2, \ldots)$.

Напишем разложение в ряд Тейлора по степеням h значения

$$\tilde{y}_{i+1}^* = \tilde{y}(x_i + h, y_0, y_1, \ldots, y_i),$$

найденное в результате применения данного приближенного метода к точным значениям $y_0, y_1, \ldots, y_i$. Будем иметь

$$\tilde{y}_{i+1}^* = b_0^{(i)} + b_1^{(i)} h + b_2^{(i)} h^2 + \ldots$$

Если для всех i выполнены равенства $b_p^{(i)} = a_p^{(i)}$ при $p = 0, 1, \ldots, m$, причем $b_{m+1}^{(i)} \neq a_{m+1}^{(i)}$ для некоторых i, то число m называется порядком точности приближенного метода. При этом погрешность метода на каждом шаге имеет порядок $O(h^{m+1})$.

Если число m определено, то для грубой оценки погрешности метода можно применить так называемый принцип Рунге [18].

Пусть, например, на каждом шаге h допущена погрешность, приблизительно пропорциональная h^{m+1} $(m \geqslant 1)$, и $2n$ представляет собой общее число шагов вычисления.

В таком случае, предполагая, что погрешность на каждом шаге одна и та же, равная Ah^{m+1}, приближенно получаем

$$\tilde{y}^*_{2n} - y_{2n} = 2nAh^{m+1}, \tag{7}$$

где A — неизвестный числовой множитель.

Согласно Рунге производим тем же методом вторичный пересчет искомого решения y с двойным шагом $H = 2h$. Тогда в силу нашего предположения будет допущена погрешность

$$\tilde{Y}^*_n - y_{2n} = nA(2h)^{m+1}, \tag{8}$$

где $\tilde{Y}^*_i$ $(i = 0, 1, 2, \ldots, n)$ — соответствующие значения в точках $\xi_i = x_0 + Hi$ результата, полученного применением данного приближенного метода с шагом $H = 2h$.

Из формул (7) и (8) выводим

$$\tilde{y}^*_{2n} - 2nAh^{m+1} = \tilde{Y}^*_n - 2^{m+1}nAh^{m+1}.$$

Отсюда находим неизвестную постоянную

$$A = \frac{\tilde{Y}^*_n - \tilde{y}^*_{2n}}{2n(2^m - 1)h^{m+1}},$$

и, следовательно,

$$y_{2n} = \tilde{y}^*_{2n} - \frac{\tilde{Y}^*_n - \tilde{y}^*_{2n}}{2^m - 1}. \tag{9}$$

Таким образом, приближенно можно положить

$$E^*_{2n} = \frac{|\tilde{Y}^*_n - \tilde{y}^*_{2n}|}{2^m - 1}. \tag{10}$$

Заметим, что формулу (9), пренебрегая текущей погрешностью, иногда используют для оценки полной погрешности решения, приближенно полагая

$$y_{2n} = \tilde{y}_{2n} - \frac{\tilde{Y}_n - \tilde{y}_{2n}}{2^m - 1}, \tag{11}$$

где $\tilde{Y}_n = \tilde{y}(x_{2n}, \tilde{Y}_0, \tilde{Y}_1, \ldots, \tilde{Y}_{n-1})$.

Применим принцип Рунге для оценки погрешности методов Рунге—Кутта и Адамса, причем ограничимся лишь рассмотрением дифференциального уравнения первого порядка

$$y' = f(x, y) \tag{12}$$

при начальном условии

$$y(x_0)=y_0. \tag{13}$$

1. Метод Рунге—Кутта. Порядок точности метода Рунге—Кутта для уравнения (1) есть $m=4$ [20], [18] (см. § 7). Поэтому из формулы (11), при сохранении введенных выше обозначений, имеем

$$y_{2n}=\tilde{y}_{2n}-\frac{1}{15}(\tilde{Y}_n-\tilde{y}_{2n}). \tag{14}$$

2. Метод Адамса. Применяя вторую интерполяционную формулу Ньютона с остаточным членом (см. [23]), получим формулу погрешности метода

$$E_i^*=\left|\int\limits_{x_i}^{x_i+h}\frac{q(q+1)(q+2)(q+3)}{4!}h^4f^{\mathrm{IV}}(\xi)\,dx\right|\leqslant$$

$$\leqslant\max|f^{\mathrm{IV}}(x)|\frac{h^5}{24}\int\limits_0^1(q^4+6q^3+11q^2+6q)\,dq=\frac{251}{720}h^5\max|f^{\mathrm{IV}}(x)|. \tag{15}$$

Приближенно можно положить

$$\frac{251}{750}\approx\frac{1}{3}\quad\text{и}\quad h^5f^{\mathrm{IV}}(x)=\Delta^4(hy').$$

Следовательно,

$$E_i^*=\frac{1}{3}\max|\Delta^4(hy')|. \tag{16}$$

Из формулы (15) вытекает, что порядок точности метода Адамса $m=4$. Кроме того, если $\Delta^4(hy')$ почти постоянны, то можно пользоваться формулой (10). Более точная оценка погрешности метода Адамса имеется у Коллатца [18].

Что касается текущей погрешности метода ε_i, то оценка ее в общем случае затруднительна, и этим вопросом здесь заниматься не будем. Заметим, однако, что различают *устойчивые схемы вычислений*, когда небольшие начальные отклонения затухают в процессе решения, и *неустойчивые схемы вычислений*, при которых даже ничтожно малые начальные отклонения неограниченно возрастают с увеличением числа шагов.

Для дифференциального уравнения (12) можно указать еще один способ заключительной проверки. Имеем

$$y(x_n)=y_0+\int\limits_{x_0}^{x_n}f[x,\ y(x)]\,dx. \tag{17}$$

Используя найденные приближенные значения

$$\tilde{y}(x_k)\quad(k=0,\ 1,\ 2,\ \ldots,\ n)$$

и применяя к правой части равенства (17) одну из квадратурных формул, получим

$$\bar{y}(x_n) = y_0 + \sum_{k=0}^{n} A_k f[x_k, \tilde{y}(x_k)],$$

где A_k — некоторые постоянные коэффициенты.

За предельную абсолютную погрешность приближенного решения $\tilde{y}(x)$ при $x_0 \leqslant x \leqslant x_n$ можно принять величину $E_n = |\tilde{y}(x_n) - \bar{y}(x_n)|$.

ЛИТЕРАТУРА К ГЛАВЕ III

[1] Степанов В. В., Курс дифференциальных уравнений, изд. 8, Физматгиз, 1959.

[2] Петровский И. Г., Лекции по теории обыкновенных дифференциальных уравнений, изд. 5, «Наука», 1964.

[3] Демидович Б. П., Дифференциальные уравнения, Арт. инж. акад., 1955.

[4] Толстов Г. П., Курс математического анализа, т. I, изд. 2, 1957, т. II, 1957.

[5] Эльсгольц Л. Э., Обыкновенные дифференциальные уравнения, Гостехиздат, 1954.

[6] Стокер Дж., Нелинейные колебания в механических и электрических системах, ИЛ, 1952, гл. III.

[7] Сикорский Ю. С., Обыкновенные дифференциальные уравнения, Гостехиздат, 1940, гл. III.

[8] Вентцель Д. А. и Шапиро Я. М., Внешняя баллистика, Оборонгиз, 1939, гл. III и VI.

[9] Биркгоф Дж. Д., Динамические системы, Гостехиздат, 1941, гл. IX.

[10] Крылов А. Н., Лекции о приближенных вычислениях, Гостехиздат, 1954, гл. VII.

[11] Гурса Э., Курс математического анализа, т. II, ч. 2, ГТТИ, 1933, гл. XIX.

[12] Гогейзель Г., Обыкновенные дифференциальные уравнения, ОНТИ, 1937, гл. II.

[13] Пиаджио Г., Интегрирование дифференциальных уравнений, ГТТИ, 1933, гл. IX.

[14] Милн В. Э., Численное решение дифференциальных уравнений, ИЛ, 1955, гл. III.

[15] Чаплыгин С. А., Новый метод приближенного интегрирования дифференциальных уравнений, Гостехиздат, 1950.

[16] Лузин Н. Н., О методе приближенного интегрирования акад. С. А. Чаплыгина, Труды ЦАГИ, 1932.

[17] Янке Е. и Эмде Ф., Таблицы функций с формулами и кривыми, Физматгиз, 1959.

[18] Коллатц Л., Численные методы решения дифференциальных уравнений, ИЛ, 1953, гл. I.

[19] Скарборо Дж., Численные методы математического анализа, ГТТИ, 1934, гл. XI, XIII.

[20] Безикович Я. С., Приближенные вычисления, Гостехиздат, 1949, гл. X.

[21] Фролов С. В., Приближенные вычисления, МВТУ, 1948.

[22] Фихтенгольц Г. М., Математика для инженеров, ч. II, вып. 2, ГТТИ, 1933, гл. XVII.
[23] Демидович Б. П., Марон И. А., Основы вычислительной математики, изд. 3, «Наука», 1966, гл. XIV.
[24] Березин И. С., Жидков Н. П., Методы вычислений, Физматгиз, 1960, т. II, гл. IX.
[25] Положий Г. Н. и др., Математический практикум, Физматгиз, 1960, гл. V.
[26] Система стандартных подпрограмм, под ред. М. Р. Шура-Бура, Физматгиз, 1958, гл. VII.
[27] Вентцель Д. А., Вентцель Е. С., Элементы теории приближенных вычислений, изд. ВИА им. Жуковского, 1949, гл. VIII.
[28] Оппоков Г. В., Численное интегрирование дифференциальных уравнений, ГТТИ, 1932, гл. II.
[29] Михлин С. Г., Смолицкий Х. Л., Приближенные методы решения дифференциальных и интегральных уравнений, серия «Справочная математическая библиотека», «Наука», гл. I, М., 1965.
[30] Канторович Л. В., Функциональный анализ и прикладная математика, УМН 3, вып. 6 (28), 1948.

ГЛАВА IV

КРАЕВЫЕ ЗАДАЧИ ДЛЯ ОБЫКНОВЕННЫХ ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ

§ 1. Общая постановка краевой задачи

Рассмотрим дифференциальное уравнение n-го порядка ($n \geqslant 2$)

$$F(x, y, y', \ldots, y^{(n)}) = 0. \tag{1}$$

Краевая задача для уравнения (1) заключается в следующем: найти решение $y = y(x)$ уравнения (1), для которого значения его производных

$$y_i^{(s)} = y^{(s)}(x_i) \quad (s = 0, 1, \ldots, \sigma_i)$$

в заданной системе точек $x = x_i$ ($i = 1, 2, \ldots, k$; $k \geqslant 2$) удовлетворяют n независимым между собой *краевым условиям*, в общем случае нелинейным:

$$V_\nu(y_1, y_1', \ldots, y_1^{(\sigma_{1\nu})}, \ldots, y_k, y_k', \ldots, y_k^{(\sigma_{k\nu})}) = 0 \tag{2}$$
$$(\nu = 1, 2, \ldots, n).$$

Так как в силу уравнения (1) производные $y^{(s)}$ порядка n и выше могут быть в общем случае выражены через искомую функцию y и ее младшие производные $y', y'', \ldots, y^{(n-1)}$, то можно считать, что

$$\sigma_{i\nu} \leqslant n - 1 \quad (i = 1, 2, \ldots, k;\ \nu = 1, 2, \ldots, n). \tag{3}$$

Краевая задача (1)—(2) часто встречается в приложениях. Приведем примеры конкретных краевых задач.

Пример 1. *Простейшая двухточечная краевая задача.* Найти функцию $y = y(x)$, удовлетворяющую дифференциальному уравнению второго порядка

$$y'' = f(x, y, y') \tag{4}$$

и принимающую при $x = a$ и $x = b$ ($a < b$) заданные значения

$$y(a) = A;\ y(b) = B.$$

Геометрически это означает, что требуется найти интегральную кривую дифференциального уравнения (4), проходящую через данные точки $M(a, A)$ и $N(b, B)$ (рис. 39).

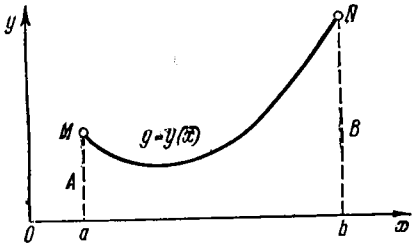

Рис. 39.

Пример 2. Видоизменением задачи, приведенной в примере 1, будет: найти такое решение $y=y(x)$ дифференциального уравнения (4), чтобы

$$y'(a)=A_1, \quad y'(b)=B_1.$$

Геометрически эта задача сводится к отысканию интегральной кривой уравнения (4), пересекающей прямые $x=a$ и $x=b$ под заданными соответственно углами α и β такими, что

$$\operatorname{tg}\alpha=A_1 \quad \text{и} \quad \operatorname{tg}\beta=B_1$$

(рис. 40).

Пример 3. Можно рассмотреть также смешанную краевую задачу: найти решение $y=y(x)$ дифференциального уравнения (4), удовлетворяющее условиям

$$y(a)=A, \quad y'(b)=B_1.$$

Иными словами, требуется найти интегральную кривую уравнения (4), проходящую через заданную точку $M(a, A)$ и пересекающую прямую $x=b$ под данным углом β, где $\operatorname{tg}\beta=B_1$ (рис. 41).

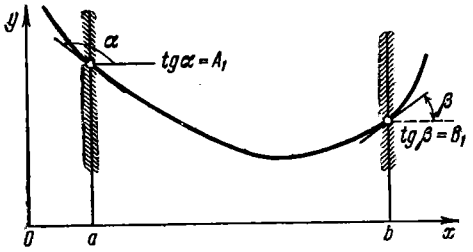

Рис. 40.

Заметим, что общая краевая задача (1)—(2) может

а) не иметь решений;

б) иметь единственное решение;

в) иметь несколько и даже бесконечно много решений.

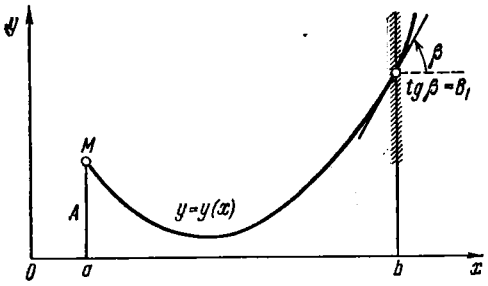

Рис. 41.

Пример 4. Краевая задача

$$y''+y=0, \quad y(0)=y(\pi)=0 \quad (5)$$

имеет бесконечно много решений вида $y=c\sin x$, где c — произвольная постоянная.

Краевая задача $y''+y=0$, $y(0)=0$, $y(b)=1$ при $0<b<\pi$ имеет единственное решение $y_b=\dfrac{\sin x}{\sin b}$, а при $b=\pi$ совсем не имеет решений (рис. 42).

В дальнейшем, как правило, будем предполагать, что решение краевой задачи существует и оно единственно. Наша цель будет заключаться в нахождении этого решения.

Специфическую особенность имеют краевые задачи (1)—(2), где одна или две из абсцисс x_i принимают значения $\pm\infty$. С такими задачами приходится сталкиваться, например, в квантовой механике. Приведем типичную постановку одной из таких задач.

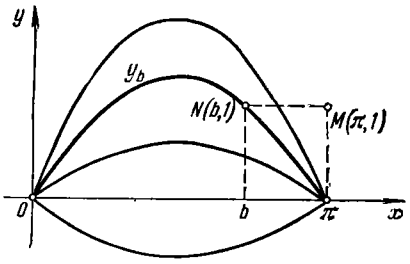

Рис. 42.

Пример 5. Найти решение $x=x(t)$ нелинейного дифференциального уравнения

$$\ddot{x}+\dot{x}f(x,\dot{x})+\varphi(x)=0 \quad \left(\dot{x}=\frac{dx}{dt}\right), \tag{6}$$

удовлетворяющее краевым условиям $x(0)=x_0$, $x(+\infty)=A$, где $A=A(x_0)$.

Таким образом, речь идет о разыскании интегральной кривой дифференциального уравнения (6), проходящей через данную точку $M(0, x_0)$, ограниченной на бесконечном интервале $(0, +\infty)$ и имеющей при $t\to+\infty$

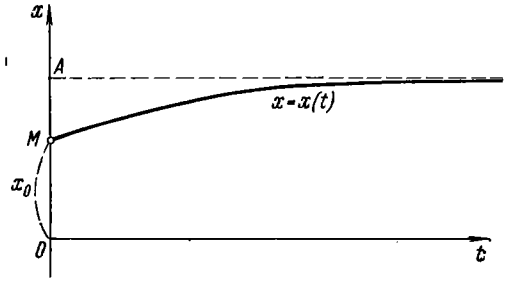

Рис. 43.

некоторую горизонтальную асимптоту $x=A$, причем величина A также должна быть определена по заданному значению x_0 (рис. 43).

Аналогично ставятся краевые задачи для систем дифференциальных уравнений.

Заметим, что краевые условия вида (2) не исчерпывают всех краевых задач. В некоторых случаях отдельные значения абсцисс x_i, в которых задаются значения искомой функции y или ее производных, также остаются неизвестными и должны быть найдены из условий задачи. К числу таких задач относится, например, *задача о поражении цели баллистическим снарядом*.

Пример 6. Как известно, дифференциальные уравнения движения снаряда с учетом сопротивления воздуха имеют вид

$$\left.\begin{aligned} \ddot{x} &= -E\cos\Theta, \\ \ddot{y} &= -E\sin\Theta - g \end{aligned}\right\} \tag{7}$$

(точки обозначают дифференцирование по времени t), где $E=E(y, v)$ — известная функция высоты y и скорости $v=\sqrt{\dot{x}^2+\dot{y}^2}$; $g=g(y)$ — ускорение силы тяжести; $\Theta=\operatorname{arctg}\frac{\dot{y}}{\dot{x}}$ — угол наклона к горизонту касательной к траектории снаряда.

Предполагая, что в момент $t=t_0$ снаряд был брошен в точке $O(0, 0)$ с начальной скоростью v_0, направленной под углом Θ_0 к горизонту, и в

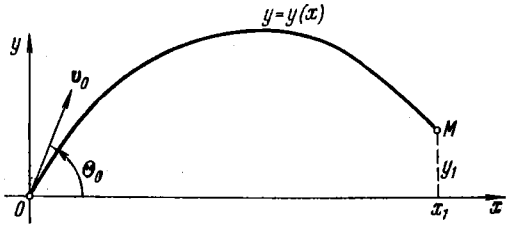

Рис. 44.

момент времени $t=t_1$ поразил неподвижную цель M с координатами $x=x_1$ и $y=y_1$ (рис. 44), получаем следующую систему краевых условий:

$$\left.\begin{aligned} &x=0,\ y=0,\ \dot{x}=v_0\cos\Theta_0,\ \dot{y}=v_0\sin\Theta_0 && \text{при } t=t_0; \\ &x=x_1,\ y=y_1 && \text{при } t=t_1; \end{aligned}\right\} \tag{8}$$

здесь угол бросания Θ_0 (параметр) и момент поражения цели t_1 также являются неизвестными.

Решив краевую задачу (7)—(8), получаем возможность найти начальный угол бросания

$$\Theta_0=\operatorname{arctg}\frac{\dot{y}_0}{\dot{x}_0},$$

где $\dot{x}_0=\dot{x}(t_0)$ и $\dot{y}_0=\dot{y}(t_0)$, при котором достигается поражение цели. Задача усложняется, если цель подвижна.

§ 2. Линейная краевая задача

Рассмотрим более подробно важный частный случай, когда дифференциальное уравнение и краевые условия линейны. Такая краевая задача называется *линейной краевой задачей*.

Линейное дифференциальное уравнение n-го порядка сокращенно можно записать в виде

$$L[y]=f(x), \tag{1}$$

где $L[y]=p_0(x)y^{(n)}+p_1(x)y^{(n-1)}+\ldots+p_n(x)y$, причем обычно предполагается, что $p_i(x)$ $(i=0, 1, \ldots, n)$ и $f(x)$ — известные непрерывные функции на данном отрезке $[a, b]$.

Для простоты будем предполагать, что в краевые условия входят две абсциссы $x_1=a$ и $x_2=b$ $(a<b)$ — концы отрезка $[a, b]$. Такие краевые условия называются *двухточечными*. Краевые условия называются линейными, если они имеют вид

$$R_\nu[y]=\gamma_\nu \quad (\nu=1, 2, \ldots, n), \tag{2}$$

где

$$R_\nu[y]=\sum_{k=0}^{n-1}\left[\alpha_k^{(\nu)}y^{(k)}(a)+\beta_k^{(\nu)}y^{(k)}(b)\right]$$

и $\alpha_k^{(\nu)}$, $\beta_k^{(\nu)}$, γ_ν — заданные постоянные, причем

$$\sum_{k=0}^{n-1}\left(|\alpha_k^{(\nu)}|+|\beta_k^{(\nu)}|\right)\neq 0 \quad \text{при} \quad \nu=1, 2, \ldots, n.$$

Например, краевые условия, приведенные в § 1 (примеры 1, 2, 3), линейны, так как их можно записать в виде

$$\alpha_0 y(a)+\alpha_1 y'(a)=\gamma_1, \quad \beta_0 y(b)+\beta_1 y'(b)=\gamma_2,$$

где α_0, α_1, β_0, β_1, γ_1, γ_2 — заданные постоянные. Действительно, для задачи примера 1 имеем

$$\alpha_0=1;\ \alpha_1=0;\ \gamma_1=A;\ \beta_0=1;\ \beta_1=0;\ \gamma_2=B \text{ и т. д.}$$

Линейными краевыми условиями являются также условия периодичности, которые в случае дифференциального уравнения второго порядка имеют вид

$$y(a)=y(b), \quad y'(a)=y'(b).$$

Линейная краевая задача состоит в нахождении функции $y=y(x)$, удовлетворяющей дифференциальному уравнению (1) и краевым условиям (2), причем последние предполагаются линейно независимыми. Линейная краевая задача называется *однородной*, если, во-первых, $f(x)\equiv 0$ при $a\leqslant x\leqslant b$, т. е. дифференциальное уравнение (1) однородно, и, во-вторых, $\gamma_\nu=0$ $(\nu=1, 2, \ldots, n)$, т. е. имеют место однородные краевые условия; в противном случае краевая задача (1)—(2) называется *неоднородной*.

Пример 1. Рассмотрим задачу об изгибе горизонтальной балки длиной l, лежащей на двух опорах $x=0$ и $x=l$, под действием распределенной поперечной нагрузки с линейной плотностью $q=q(x)$ (рис. 45).

Из курса сопротивления материалов известно, что вертикальный прогиб однородной балки приближенно удовлетворяет линейному дифференциальному уравнению

$$[EI(x)y'']''=q(x), \tag{3}$$

где $EI(x)$ — жесткость балки при изгибе, причем изгибающий момент M и поперечная сила Q определяются из соотношений

$$M = EI(x)\,y'' \quad \text{и} \quad Q = M' = [EI(x)\,y'']'.$$

Краевые условия зависят от способа заделки концов балки. Приведем основные случаи.

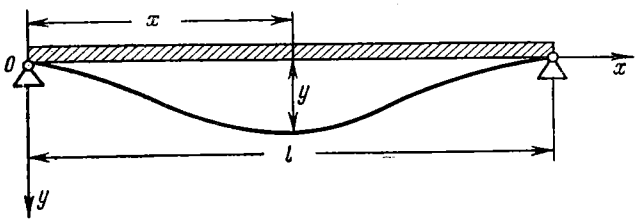

Рис. 45.

1. Конец свободен. Нулю равны изгибающий момент M и поперечная сила Q. Поэтому *краевые условия для свободного конца* балки суть

$$y'' = 0 \quad \text{и} \quad y''' = 0. \tag{4}$$

2. Конец шарнирно оперт. Нулю равны прогиб y и изгибающий момент M. Поэтому *краевые условия для шарнирно опертого конца* суть

$$y = 0 \quad \text{и} \quad y'' = 0. \tag{4'}$$

3. Конец жестко заделан. Нулю равны прогиб y и угол поворота $\varphi = \operatorname{arctg} y'$. Поэтому *краевые условия жестко заделанного конца* суть

$$y = 0 \quad \text{и} \quad y' = 0. \tag{4''}$$

Возможны также другие, более сложные случаи краевых условий.

Задача (3)—(4), очевидно, является линейной краевой задачей.

Пример 2. Пусть жесткость балки EI постоянна, тогда уравнение (3) для прогиба y заменяется следующим уравнением:

$$EIy^{\mathrm{IV}} = q(x). \tag{5}$$

Предположим, что балка шарнирно закреплена на конце $x=0$ и жестко заделана на конце $x=l$. В таком случае для прогиба y выполнены краевые (граничные) условия:

$$\left.\begin{aligned} y(0) = 0, \quad y''(0) = 0, \\ y(l) = 0, \quad y'(l) = 0. \end{aligned}\right\} \tag{6}$$

Краевые условия (6) являются линейными однородными, так как их можно записать в виде

$$\left.\begin{aligned} 1y(0) + 0y''(0) + 0y(l) + 0y'(l) = 0, \\ 0y(0) + 1y''(0) + 0y(l) + 0y'(l) = 0, \\ 0y(0) + 0y''(0) + 1y(l) + 0y'(l) = 0, \\ 0y(0) + 0y''(0) + 0y(l) + 1y'(l) = 0. \end{aligned}\right\}$$

Краевую задачу (5)—(6) решить нетрудно. Предполагая для простоты, что плотность нагрузки постоянна: $q(x) = p$, будем иметь

$$EIy = \frac{px^4}{24} + c_1x^3 + c_2x^2 + c_3x + c_4.$$

Из граничных условий (6) вытекает

$$c_1 = -\frac{pl}{16}, \quad c_2 = 0, \quad c_3 = \frac{pl^3}{48}, \quad c_4 = 0.$$

Таким образом, искомое решение есть

$$y = \frac{p}{48EI}(2x^4 - 3lx^3 + l^3x).$$

Этот пример показывает, что в случае, когда можно найти общее решение дифференциального уравнения, двухточечная краевая задача не более трудна, чем задача с начальными условиями. Однако если общее решение уравнения не может быть найдено регулярным путем, то решение краевой задачи приводит к новым трудностям, так как не имеется начальной точки, исходя из которой можно было бы построить решение одним из рассмотренных выше методов.

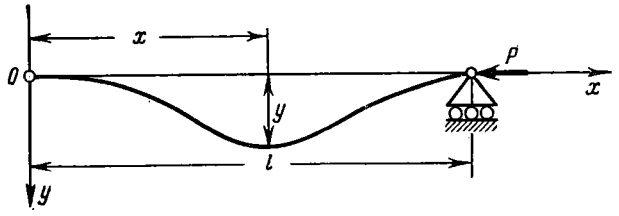

Рис. 46.

Однородная краевая задача (1) — (2) всегда имеет тривиальное решение $y(x) \equiv 0$. Однако во многих случаях представляют интерес нетривиальные решения этой задачи, которые существуют не всегда. Поэтому в дифференциальное уравнение (1) или в краевые условия (2) вводят параметр λ, варьируя который можно добиться, чтобы при некоторых его значениях соответствующая краевая задача имела нетривиальные решения. Эти исключительные значения параметра называются *собственными значениями* или *характеристическими числами* задачи, а отвечающие им нетривиальные решения — *собственными функциями* задачи. Таким образом, приходим к так называемой *задаче о собственных значениях* — важнейшей задаче современного математического анализа.

Пример 3. Рассмотрим задачу о продольном изгибе стержня постоянной жесткости EI под действием сжимающей силы P, направленной вдоль оси стержня *(задача Эйлера)*. Предположим, что левый конец стержня $x=0$ заделан, а правый $x=l$ оперт (рис. 46). Как известно, величина $y=y(x)$ — отклонение стержня от его оси — удовлетворяет дифференциальному уравнению

$$y^{\mathrm{IV}} + \frac{P}{EI}y'' = 0, \tag{7}$$

где P играет роль параметра. Кроме того, согласно способу заделки стержня на его концах должны быть выполнены краевые условия:

$$y(0)=y'(0)=0, \quad y(l)=y''(l)=0. \tag{8}$$

Требуется найти минимальное значение силы P — так называемую *критическую силу*, при которой возможен продольный изгиб. Математически этот вопрос сводится к определению наименьшего положительного значения параметра P, при котором краевая задача (7)—(8) имеет нетривиальное решение.

Общее решение дифференциального уравнения (7) имеет вид

$$y=c_1+c_2x+c_3\cos\alpha x+c_4\sin\alpha x,$$

где

$$\alpha=\sqrt{\frac{P}{EI}} \tag{9}$$

и c_1, c_2, c_3, c_4 — произвольные постоянные.

На основании первых двух краевых условий (8) получаем

$$0=c_1+c_3, \quad 0=c_2+c_4\alpha.$$

Отсюда $c_3=-c_1$, $c_4=-\frac{c_2}{\alpha}$, и, следовательно,

$$y=c_1(1-\cos\alpha x)+c_2\left(x-\frac{\sin\alpha x}{\alpha}\right).$$

Два вторых краевых условия (8) в связи с тем, что $\alpha\neq 0$, дают

$$\left.\begin{aligned} 0&=c_1(1-\cos\alpha l)+c_2\left(l-\frac{\sin\alpha l}{\alpha}\right),\\ 0&=c_1\cos\alpha l+c_2\frac{\sin\alpha l}{\alpha}.\end{aligned}\right\} \tag{10}$$

Таким образом, для определения постоянных c_1 и c_2 получена однородная линейная система. Согласно смыслу задачи представляют интерес лишь решения, отличные от нуля; поэтому определитель системы (10) должен быть равен нулю, т. е.

$$\begin{vmatrix} 1-\cos\alpha l & l-\dfrac{\sin\alpha l}{\alpha}\\ \cos\alpha l & \dfrac{\sin\alpha l}{\alpha}\end{vmatrix}=0 \tag{11}$$

или

$$\sin\alpha l-\alpha l\cos\alpha l=0.$$

Отсюда будем иметь $\operatorname{tg}\alpha=\alpha l$, и, следовательно,

$$\alpha l=\mu, \tag{12}$$

где $\mu=4{,}493\ldots$ — наименьший положительный корень трансцендентного уравнения $\operatorname{tg}\mu=\mu$.

Из формул (9) и (12) выводим

$$P_{\text{крит}}=\frac{\mu^2EI}{l^2}\approx 20{,}187\frac{EI}{l^2}.$$

Точное решение краевой задачи возможно в редких случаях. Поэтому в дальнейшем будут рассмотрены приближенные методы решения краевых задач, причем для простоты в основном ограничимся линейным дифференциальным уравнением второго порядка и линейными краевыми условиями простейшего вида — *точечно разделенными* *).

§ 3. Редукция к задаче Коши двухточечной краевой задачи для линейного уравнения второго порядка

Дано линейное дифференциальное уравнение

$$y'' + p(x)\,y' + q(x)\,y = f(x), \tag{1}$$

где функции $p(x)$, $q(x)$, $f(x)$ непрерывны, и требуется найти его решение, удовлетворяющее краевым условиям:

$$\alpha_0 y(a) + \alpha_1 y'(a) = A, \quad \beta_0 y(b) + \beta_1 y'(b) = B \tag{2}$$

(α_0, α_1, β_0, β_1, A, B — заданные постоянные, причем $|\alpha_0| + |\alpha_1| \neq 0$ и $|\beta_0| + |\beta_1| \neq 0$).

Решение будем искать в виде линейной комбинации

$$y = cu + v, \tag{3}$$

где $u = u(x)$ — ненулевое решение соответствующего однородного уравнения

$$u'' + p(x)\,u' + q(x)\,u = 0, \tag{4}$$

а $v = v(x)$ — некоторое решение данного неоднородного уравнения (1)

$$v'' + p(x)\,v' + q(x)\,v = f(x). \tag{5}$$

Очевидно, функция y, определяемая формулой (3), где c произвольно, есть решение уравнения (1).

Потребуем, чтобы первое краевое условие (2) выполнялось для функции y при любом c. Используя это краевое условие, будем иметь

$$c\alpha_0 u(a) + \alpha_0 v(a) + c\alpha_1 u'(a) + \alpha_1 v'(a) = A$$

или

$$c\,[\alpha_0 u(a) + \alpha_1 u'(a)] + \alpha_0 v(a) + \alpha_1 v'(a) = A. \tag{6}$$

Для того чтобы равенство (6) было справедливо при любом c, необходимо и достаточно, чтобы коэффициент при c обращался в нуль, т. е. должны быть выполнены равенства

$$\alpha_0 u(a) + \alpha_1 u'(a) = 0, \tag{7}$$

$$\alpha_0 v(a) + \alpha_1 v'(a) = A. \tag{8}$$

Для обеспечения равенств (7), (8) достаточно, например, положить

$$u(a) = \alpha_1 k, \quad u'(a) = -\alpha_0 k, \tag{9}$$

*) Каждое из этих условий будет содержать лишь одну из абсцисс.

где постоянная k отлична от нуля;

$$v(a)=\frac{A}{\alpha_0},\quad v'(a)=0, \tag{10}$$

если $\alpha_0 \neq 0$, и

$$v(a)=0,\quad v'(a)=\frac{A}{\alpha_1}, \tag{11}$$

если $\alpha_1 \neq 0$.

Отсюда видно, что u есть решение задачи Коши для однородного уравнения (4), удовлетворяющее начальным условиям (9), а v есть решение задачи Коши для неоднородного уравнения (5), удовлетворяющее начальным условиям (10) или (11). При этом для любого c функция $y=cu+v$ удовлетворяет краевому условию на конце $x=a$.

Подберем теперь постоянную c так, чтобы функция y удовлетворяла краевому условию (2) на конце $x=b$. Это дает

$$c[\beta_0 u(b)+\beta_1 u'(b)]+[\beta_0 v(b)+\beta_1 v'(b)]=B,$$

откуда

$$c=\frac{B-[\beta_0 v(b)+\beta_1 v'(b)]}{\beta_0 u(b)+\beta_1 u'(b)}.$$

При этом предполагается, что знаменатель

$$\beta_0 u(b)+\beta_1 u'(b)\neq 0. \tag{12}$$

Таким образом, краевая задача (1)—(2) сведена к двум задачам Коши для функций $u(x)$ и $v(x)$.

Заметим, что если обеспечено условие (12), то краевая задача (1)—(2) имеет единственное решение. В противном случае она или совсем не имеет решений, или их бесчисленное множество.

Замечание. Если исходное уравнение (1) однородное, т. е. $f(x)\equiv 0$ и, кроме того, $A=0$, то в силу условий (10) или (11) имеем $v(a)=0$ и $v'(a)=0$ и, следовательно, $v\equiv 0$. Поэтому $y=cu(x)$, где $u(x)$ есть решение уравнения (4), удовлетворяющее начальным условиям (9). В этом случае

$$c=\frac{B}{\beta_0 u(b)+\beta_1 u'(b)}.$$

Пример 1. Найти решение однородного уравнения

$$y''+y\operatorname{ch}x=0, \tag{13}$$

удовлетворяющее краевым условиям

$$y(0)=0,\quad y(1)=1. \tag{14}$$

Решение. Сравнивая условия (14) с общими краевыми условиями (2), видим, что $\alpha_0=1$, $\alpha_1=0$, $A=0$; $\beta_0=1$, $\beta_1=0$, $B=1$. В силу приведенного выше замечания решение ищем в виде

$$y=cu(x),$$

где $u(x)$ — решение однородного уравнения

$$u'' + u \operatorname{ch} x = 0, \tag{15}$$

удовлетворяющее начальным условиям (9), где принято $k = -1$, т. е.

$$u(0) = 0, \quad u'(0) = 1. \tag{16}$$

Из второго краевого условия (14) получаем $cu(1) = 1$, отсюда $c = \frac{1}{u(1)}$.

Решая любым численным методом задачу Коши (15)—(16), находим $u(x)$, а следовательно, и постоянную c, после чего определяем y.

Таблица 60

Решение краевой задачи для однородного уравнения (13)

x	0	0,2	0,4	0,6	0,8	1
$u(x)$	0	0,199	0,389	0,563	0,710	0,819
$y(x)$	0	0,243	0,475	0,687	0,867	1

В данном случае соответствующая задача Коши решена в гл. III, § 12. Необходимые данные находятся в таблице 58. Так как $u(1) = 0{,}819$, то $c = 1{,}221$ и $y = 1{,}221\, u(x)$. Окончательные результаты вычислений приведены в таблице 60.

§ 4. Метод конечных разностей

Рассмотрим линейное дифференциальное уравнение

$$y'' + p(x)\, y' + q(x)\, y = f(x) \tag{1}$$

с двухточечными линейными краевыми условиями

$$\left.\begin{aligned} \alpha_0 y(a) + \alpha_1 y'(a) &= A, \\ \beta_0 y(b) + \beta_1 y'(b) &= B \end{aligned}\right\} \tag{2}$$

$$(|\alpha_0| + |\alpha_1| \neq 0, \quad |\beta_0| + |\beta_1| \neq 0),$$

где $p(x)$, $q(x)$ и $f(x)$ непрерывны на отрезке $[a, b]$.

Одним из наиболее простых методов решения этой краевой задачи является сведение ее к системе конечно-разностных уравнений. Для этого разобьем основной отрезок $[a, b]$ на n равных частей длины h (шаг), где

$$h = \frac{b-a}{n}.$$

Точки разбиения имеют абсциссы:

$$x_i = x_0 + ih \quad (i = 0, 1, 2, \ldots, n), \quad x_0 = a, \quad x_n = b.$$

Значения в точках деления x_i искомой функции $y=y(x)$ (рис. 47) и ее производных $y'=y'(x)$, $y''=y''(x)$ обозначим соответственно через $y_i=y(x_i)$, $y_i'=y'(x_i)$, $y_i''=y''(x_i)$. Введем также обозначения: $p_i=p(x_i)$, $q_i=q(x_i)$, $f_i=f(x_i)$.

Заменяя производные симметричными конечно-разностными отношениями для внутренних точек x_i отрезка $[a, b]$, будем иметь

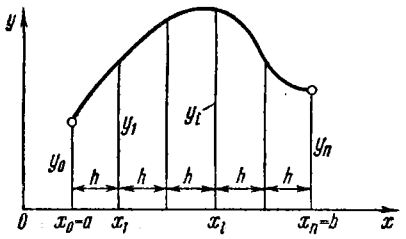

Рис. 47.

$$y_i'=\frac{y_{i+1}-y_{i-1}}{2h},$$

$$y_i''=\frac{y_{i+2}-2y_i+y_{i-1}}{h^2} \qquad (3)$$

$$(i=1,\ 2,\ \ldots,\ n-1).$$

Для концевых точек $x_0=a$ и $x_n=b$, чтобы не выходить за пределы отрезка $[a, b]$, можно положить

$$y_0'=\frac{y_1-y_0}{h},\quad y_n'=\frac{y_{n-1}-y_n}{-h}. \qquad (4)$$

Однако если функция $y=y(x)$ достаточно гладкая, то более точные значения дают формулы [2]

$$y_0'=\frac{-y_2+4y_1-3y_0}{2h} \qquad (5)$$

и

$$y_n'=\frac{3y_n-4y_{n-1}+y_{n-2}}{2h},\quad n\geqslant 2. \qquad (6)$$

Действительно, полагая, например, $y_1=y(x_0+h)$ и $y_2=y(x_0+2h)$ и используя формулу Тейлора, будем иметь

$$y_1=y_0+hy_0'+\frac{h^2}{2!}y_0''+\frac{h^3}{3!}y_0'''+\ldots,$$

$$y_2=y_0+2hy_0'+\frac{(2h)^2}{2!}y_0''+\frac{(2h)^3}{3!}y_0'''+\ldots$$

Отсюда

$$\frac{-y_2+4y_1-3y_0}{2h}=y_0'+O(h^2),$$

где через $O(h^2)$, как обычно, обозначена величина порядка h^2. Аналогично показывается, что

$$\frac{3y_n-4y_{n-1}+y_{n-2}}{2h}=y_n'+O(h^2).$$

Используя формулы (3), дифференциальное уравнение (1) во

внутренних точках $x=x_i$ $(i=1, 2, \ldots, n-1)$ приближенно можно заменить линейной системой уравнений

$$\frac{y_{i+1}-2y_i+y_{i-1}}{h^2}+p_i\frac{y_{i+1}-y_{i-1}}{2h}+q_iy_i=f_i \tag{7}$$
$$(i=1, 2, \ldots, n-1).$$

Кроме того, в силу формул (5) и (6) краевые условия (2) дополнительно дают еще два уравнения:

$$\left.\begin{aligned}\alpha_0y_0+\alpha_1\frac{-y_2+4y_1-3y_0}{2h}&=A,\\ \beta_0y_n+\beta_1\frac{3y_n-4y_{n-1}+y_{n-2}}{2h}&=B.\end{aligned}\right\} \tag{8}$$

Таким образом получена линейная система $n+1$ уравнений с $n+1$ неизвестными $y_0, y_1, y_2, \ldots, y_n$, представляющими собой значения искомой функции $y=y(x)$ в точках $x_0, x_1, x_2, \ldots, x_n$. Решив эту систему, если это возможно, получим таблицу значений искомой функции y.

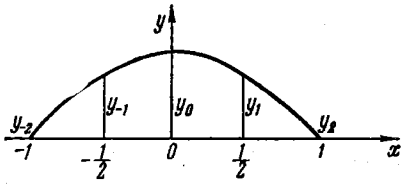

Рис. 48.

Пример 1. Методом конечных разностей найти решение краевой задачи

$$\left.\begin{aligned}y''+(1+x^2)y&=-1,\\ y(-1)=y(1)&=0.\end{aligned}\right\} \tag{9}$$

Механически уравнения (9) представляют собой дифференциальные уравнения для изгибающего момента некоторого бруса с переменным поперечным сечением и шарнирно закрепленными концами.

Для грубого решения выберем шаг $h=1/2$.

Полагая $x_{-2}=-1$, $x_{-1}=-1/2$, $x_0=0$, $x_1=1/2$, $x_2=1$, ввиду симметрии уравнения и краевых условий будем иметь $y_{-2}=y_2=0$, $y_{-1}=y_1$ (рис. 48). Таким образом, нужно определить лишь две ординаты y_0 и y_1.

Полагая $x=0$ и пользуясь симметричными формулами для производных, будем иметь

$$\frac{y_{-1}-2y_0+y_1}{1/4}+y_0=-1,$$

где $y_{-1}=y_1$. Аналогично при $x=1/2$ получаем

$$\frac{y_0-2y_1+y_2}{1/4}+\left(1+\frac{1}{4}\right)y_1=-1.$$

Следовательно, используя краевое условие $y_2=0$, имеем систему

$$-7y_0+8y_1=-1, \quad 4y_0-6\frac{3}{4}y_1=-1,$$

откуда $y_0=0{,}967$; $y_1=0{,}721$.

При большом n непосредственное решение системы (7)—(8) становится затруднительным. В этом случае решение краевой задачи

целесообразно заменить, используя результаты предыдущего параграфа, решением двух задач Коши.

Например, предполагая, что $\alpha_0 \neq 0$, имеем $y = cu + v$, где u находится, как решение задачи Коши

$$u'' + p(x)u' + q(x)u = 0, \quad u(a) = \alpha_1, \quad u'(a) = -\alpha_0, \tag{10}$$

а v — как решение задачи Коши

$$v'' + p(x)v' + q(x)v = f(x), \quad v(a) = \frac{A}{\alpha_0}, \quad v'(a) = 0. \tag{11}$$

Постоянная c в силу краевых условий (2) имеет значение

$$c = \frac{B - [\beta_0 v(b) + \beta_1 v'(b)]}{\beta_0 u(b) + \beta_1 u'(b)}. \tag{12}$$

Заменяя дифференциальное уравнение (10) соответствующим конечно-разностным уравнением, получим

$$\frac{u_{i-1} - 2u_i + u_{i+1}}{h^2} + p_i \frac{u_{i+1} - u_{i-1}}{2h} + q_i u_i = 0 \qquad (i = 1, 2, \ldots, n-1),$$

$$u_0 = \alpha_1, \quad \frac{-u_2 + 4u_1 - 3u_0}{2h} = -\alpha_0.$$

Отсюда будем иметь

$$u_0 = \alpha_1, \qquad u_1 = \frac{\alpha_1(1 + p_1 h) - \alpha_0 h\left(1 + \frac{p_1}{2}h\right)}{1 + p_1 h + \frac{q_1}{2}h^2},$$

$$u_{i+1} = \frac{(2 - q_i h^2)u_i - \left(1 - \frac{p_i}{2}h\right)u_{i-1}}{1 + \frac{p_i}{2}h}$$

$(i = 1, 2, \ldots, n-1)$. С помощью этого процесса находим

$$u(b) = u_n, \qquad u'(b) = \frac{3u_n - 4u_{n-1} + u_{n-2}}{2h}. \tag{13}$$

Аналогично, заменяя дифференциальное уравнение (11) конечно-разностным уравнением, будем иметь

$$\left.\begin{aligned} &\frac{v_{i-1} - 2v_i + v_{i+1}}{h^2} + p_i \frac{v_{i+1} - v_{i-1}}{2h} + q_i v_i = f_i \\ &\qquad (i = 1, 2, \ldots, n-1), \\ &v_0 = \frac{A}{\alpha_0}, \quad \frac{-v_2 + 4v_1 - 3v_0}{2h} = 0. \end{aligned}\right\} \tag{14}$$

Отсюда получаем

$$\left.\begin{aligned} v_0 &= \frac{A}{\alpha_0}, \quad v_1 = \frac{\frac{1}{2} f_1 h^2 + \frac{A}{\alpha_0}(1+p_1 h)}{1+p_1 h+\frac{q_1}{2}h^2}, \\ v_{i+1} &= \frac{(2-q_i h^2)\, v_i - \left(1-\frac{p_i}{2}h\right) v_{i-1} + f_i h^2}{1+\frac{p_i}{2}h} \\ &\qquad (i=1,\ 2,\ \ldots,\ n-1). \end{aligned}\right\} \tag{15}$$

Таким образом, последовательным процессом можно вычислить $v(b)=v_n$ и

$$v'(b) = \frac{3y_n - 4y_{n-1} + y_{n-2}}{2h}.$$

Тем самым мы получаем возможность из формулы (12) найти постоянную c, а затем определить значения y_i $(i=0,\ 1,\ 2,\ \ldots,\ n)$ по формуле

$$y_i = cu_i + v_i \quad (i=0,\ 1,\ 2,\ \ldots,\ n).$$

Формулы для расчета упрощаются, если краевые условия имеют вид $y(a)=A$, $y(b)=B$. В этом случае $\alpha_0=\beta_0=1$, $\alpha_1=\beta_1=0$, и мы получим

$$\left.\begin{aligned} u_0 &= 0, \quad u_1 = -\frac{h\left(1+\frac{p_1}{2}h\right)}{1+p_1 h+\frac{q_1}{2}h^2}, \\ v_0 &= A, \quad v_1 = \frac{A(1+p_1 h)+\frac{1}{2}f_1 h^2}{1+p_1 h+\frac{q_1}{2}h^2}, \end{aligned}\right\} \tag{16}$$

причем $c=(B-v_n)/u_n$.

Замечание. Приближенные значения y можно уточнить, производя вычисления с двумя различными шагами. Пусть y_h и y_H — приближенные значения y при шаге h и, соответственно, шаге H. Предполагая, что ошибка пропорциональна квадрату шага, будем иметь

$$y - y_h \approx c^* h^2, \quad y - y_H \approx c^* H^2.$$

Отсюда, считая, что постоянная c^* не зависит от шага, получим

$$\frac{y-y_h}{h^2} = \frac{y-y_H}{H^2},$$

и, следовательно,

$$y = \frac{H^2 y_h - h^2 y_H}{H^2-h^2} = y_h + \frac{h^2}{H^2-h^2}(y_h - y_H). \tag{17}$$

§ 5. Метод прогонки

При применении метода конечных разностей к краевым задачам для дифференциальных уравнений второго порядка получается «трехчленная система» линейных алгебраических уравнений, каждое из которых содержит три соседних неизвестных. Для решения такой системы разработан специальный метод, получивший название *метода прогонки*.

Рассмотрим линейное дифференциальное уравнение

$$y'' + p(x)\,y' + q(x)\,y = f(x) \tag{1}$$

с двухточечными линейными краевыми условиями

$$\alpha_0 y(a) + \alpha_1 y'(a) = A, \quad \beta_0 y(b) + \beta_1 y'(b) = B \tag{2}$$

$$(|\alpha_0| + |\alpha_1| \neq 0; \quad |\beta_0| + |\beta_1| \neq 0)$$

в предположении, что функции $p(x)$, $q(x)$, $f(x)$ непрерывны на $[a, b]$.

От дифференциального уравнения (1) обычным приемом перейдем к конечно-разностному уравнению. Для этого разобьем отрезок $[a, b]$ на n равных частей с шагом $h = \frac{b-a}{n}$. Полагая $x_i = x_0 + ih$, $x_0 = a$, $x_n = b$ $(i = 0, 1, 2, \ldots, n)$ и вводя обозначения: $p(x_i) = p_i$, $q(x_i) = q_i$, $f(x_i) = f_i$, $y(x_i) = y_i$, для внутренних точек $x = x_i$ $(i = 1, 2, \ldots, n-1)$ отрезка $[a, b]$ вместо дифференциального уравнения (1) получаем систему конечноразностных уравнений

$$\frac{y_{i+1} - 2y_i + y_{i-1}}{h^2} + p_i \frac{y_{i+1} - y_{i-1}}{2h} + q_i y_i = f_i$$

$$(i = 1, 2, \ldots, n-1).$$

Отсюда после соответствующих преобразований будем иметь

$$y_{i+1} + m_i y_i + n_i y_{i-1} = \hat{f}_i h^2 \qquad (i = 1, 2, \ldots, n-1), \tag{3}$$

где для краткости положено

$$m_i = -\frac{2 - q_i h^2}{1 + \frac{p_i}{2} h}, \quad n_i = \frac{1 - \frac{p_i}{2} h}{1 + \frac{p_i}{2} h}, \quad \hat{f}_i = \frac{f_i}{1 + \frac{p_i}{2} h}. \tag{4}$$

Для производных на концах $x_0 = a$ и $x_0 = b$ берем односторонние производные

$$y'_0 = \frac{y_1 - y_0}{h} \quad \text{и} \quad y'_n = \frac{y_{n-1} - y_n}{-h}.$$

Отсюда согласно условиям (2), получим

$$\alpha_0 y_0 + \alpha_1 \frac{y_1 - y_0}{h} = A, \qquad \beta_0 y_0 + \beta_1 \frac{y_{n-1} - y_n}{-h} = B. \tag{5}$$

Линейная система (3), (5) состоит из $n+1$ уравнений первой степени относительно неизвестных $y_0, y_1, \ldots, y_n$. Эту систему можно решить обычным способом. Однако мы сейчас укажем другой, более короткий путь, получивший название *метода прогонки*.

Разрешая уравнение (3) относительно y_i, будем иметь

$$y_i = \frac{\hat{f}_i}{m_i} h^2 - \frac{1}{m_i} y_{i+1} - \frac{n_i}{m_i} y_{i-1}. \tag{6}$$

Предположим, что с помощью полной системы (3), (5) из уравнения (6) исключена неизвестная y_{i-1}. Тогда это уравнение примет вид

$$y_i = c_i (d_i - y_{i+1}), \tag{7}$$

где c_i, d_i $(i = 1, 2, \ldots, n-1)$ — некоторые коэффициенты. Отсюда

$$y_{i-1} = c_{i-1} (d_{i-1} - y_i).$$

Подставляя это выражение в уравнение (3), получим

$$y_{i+1} + m_i y_i + n_i c_{i-1} (d_{i-1} - y_i) = \hat{f}_i h^2,$$

и, следовательно,

$$y_i = \frac{(\hat{f}_i h^2 - n_i c_{i-1} d_{i-1}) - y_{i+1}}{m_i - n_i c_{i-1}}. \tag{8}$$

Сравнивая формулы (7) и (8), получим для определения c_i и d_i рекуррентные формулы:

$$c_i = \frac{1}{m_i - n_i c_{i-1}}, \quad d_i = \hat{f}_i h^2 - n_i c_{i-1} d_{i-1} \quad (i = 1, \ldots, n-1). \tag{9}$$

Определим теперь c_0 и d_0. Из первого краевого условия (5) получаем

$$y_0 = \frac{Ah - \alpha_1 y_1}{\alpha_0 h - \alpha_1}.$$

С другой стороны, из формулы (7) при $i = 0$ имеем

$$y_0 = c_0 (d_0 - y_1). \tag{10}$$

Сравнивая последние два равенства, находим

$$c_0 = \frac{\alpha_1}{\alpha_0 h - \alpha_1}, \quad d_0 = \frac{Ah}{\alpha_1}. \tag{11}$$

На основании формул (9), (11) последовательно определяются коэффициенты c_i, d_i $(i = 1, 2, \ldots, n-1)$ до c_{n-1} и d_{n-1} включительно (прямой ход).

Обратный ход начинается с определения y_n. Используя второе краевое условие (5) и формулу (7) при $i = n-1$, получим систему двух уравнений

$$\beta_0 y_n + \beta_1 \frac{y_{n-1} - y_n}{-h} = B, \quad y_{n-1} = c_{n-1} (d_{n-1} - y_n). \tag{12}$$

Решая ее относительно y_n, будем иметь

$$y_n = \frac{Bh + \beta_1 c_{n-1} d_{n-1}}{\beta_0 h + \beta_1 (c_{n-1} + 1)}. \tag{13}$$

Теперь по формуле (7) последовательно находим $y_{n-1}, y_{n-2}, y_{n-3}, \ldots$ $\ldots, y_0$.

Для контроля можно проверить выполнение первого краевого условия. Вычисления удобно расположить в виде таблицы 61.

Таблица 61

Схема метода прогонки

i	0	1	2	...	$n-2$	$n-1$	n
c_i	c_0 (11)	c_1	c_2	...	c_{n-2}	c_{n-1}	
d_i	d_0 (11)	d_1	d_2	...	d_{n-2}	d_{n-1}	
x_i	$x_0 = a$	x_1	x_2	...	x_{n-2}	x_{n-1}	$x_n = b$
y_i	y_0	y_1	y_2	...	y_{n-2}	y_{n-1}	y_n (13)

Для простейших краевых условий $y(a) = A$, $y(b) = B$ формулы для c_0, d_0, y_0 и y_n упрощаются. А именно, полагая $\alpha_0 = 1$, $\alpha_1 = 0$ и $\beta_0 = 1$, $\beta_1 = 0$, из формул (11) будем иметь: $c_0 = 0$, $d_0 = \infty$, $c_0 d_0 = A$. Отсюда

$$c_1 = \frac{1}{m_1}, \quad d_1 = \hat{f}_1 h^2 - n_1 A; \tag{14}$$

причем

$$y_n = B, \quad y_0 = A. \tag{15}$$

Заметим, что метод прогонки обладает устойчивым вычислительным алгоритмом [3], [4], т. е. ошибки округления не вызывают неограниченного возрастания погрешности решения.

Пример 1. Методом прогонки решить краевую задачу

$$y'' = x + y; \tag{16}$$

$$y(0) = 0, \quad y(1) = 0. \tag{17}$$

Решение. Примем $h = 0{,}1$ и от уравнения (16) и краевых условий (17) перейдем к соответствующим конечно-разностным уравнениям

$$y_{i+1} + m_i y_i + n_i y_{i-1} = \hat{L}_i h^2 \quad (i = 1, 2, \ldots, n-1), \quad y_0 = 0, \quad y_{10} = 0,$$

где $m_i = -2-h^2$, $n_i = 1$, $f_i = x_i = ih$. Согласно формулам (14) имеем $c_1 = -0{,}498$, $d_1 = 0{,}001$. Формулы (9) в нашем случае дают

$$c_i = \frac{1}{-2-h^2-c_{i-1}}, \quad d_i = ih^3 - c_{i-1}d_{i-1} \quad (i=1,\ 2,\ \ldots,\ n-1).$$

Найденные значения c_i и d_i $(i=1,\ 2,\ \ldots,\ 9)$ записываем в первых двух строках таблицы 62. Затем, используя формулу (7) и известное значение $y_{10}=0$, вычисляем $y_9,\ y_8,\ \ldots,\ y_1$.

Для сравнения в последней строке таблицы 62 даны значения точного решения $\tilde{y} = \frac{2e}{e^2-1}\operatorname{sh} x - x$.

Таблица 62

Решение краевой задачи (16) — (17) методом прогонки

i	0	1	2	3	4	5
c_i	0	—0,498	—0,662	—0,878	—0,890	—0,900
d_i		0,001	0,002	0,004	0,008	0,012
y_i	0	—0,025	—0,049	—0,072	—0,078	—0,081
$\tilde{y}_i$	0	—0,015	—0,029	—0,041	—0,050	—0,057

Продолжение таблицы 62

i	6	7	8	9	10
c_i	—0,908	—0,915	—0,921	—0,926	
d_i	0,016	0,022	0,028	0,035	
y_i	—0,078	—0,070	—0,055	—0,032	0
$\tilde{y}_i$	—0,058	—0,054	—0,044	—0,026	0

Пример 2. Методом прогонки решить краевую задачу

$$y'' - 2xy' - 2y = -4x; \tag{18}$$

$$y(0) - y'(0) = 0, \qquad y(1) = 3{,}718. \tag{19}$$

Решение. Примем $h=0{,}1$ и от уравнений (18) и (19) перейдем к конечно-разностным соотношениям

$$y_{i+1}+m_i y_i+n_i y_{i-1}=\hat{f}_i h^2 \qquad (i=1,\ 2,\ \ldots,\ n-1),$$

$$y_0-\frac{y_1-y_0}{h}=0, \qquad y_n=3{,}718,$$

где

$$m_i=-\frac{2+2h^2}{1-ih^2}, \qquad n_i=\frac{1+ih^2}{1-ih^2}, \qquad \hat{f}_i=-\frac{4ih}{1-ih^2}.$$

Значения m_i, n_i и $\hat{f}_i h^2$ для $i=0,\ 1,\ 2,\ \ldots,\ 9$ записываются в таблицу 63.

Таблица 63

i	0	1	2	3	4
m_i	−2,020	−2,040	−2,060	−2,083	−2,105
n_i	1	1,020	1,040	1,062	1,084
$\hat{f}_i h^2$	0	−0,004	−0,008	−0,012	−0,017

Продолжение таблицы 63

i	5	6	7	8	9
m_i	−2,127	−2,149	−2,172	−2,196	−2,20
n_i	1,106	1,128	1,150	1,174	1,198
$\hat{f}_i h^2$	−0,021	−0,025	−0,030	−0,035	−0,040

Пользуясь этой таблицей, по формулам (11) находим значения c_0 и d_0, а затем по формулам (9) находим c_i и d_i $(i=1,\ 2,\ 3,\ \ldots,\ 9)$. Эти значения записываем в первых двух строках таблицы 64.

Используя значение $y_n=3{,}718$, начинаем обратный ход и по формулам (7), последовательно вычисляя $y_9,\ y_8,\ \ldots,\ y_0$. Полученные результаты записываем в третьей строке таблицы 64.

Для сравнения в последней строке таблицы даны значения точного решения $\tilde{y}=x+e^{x^2}$.

Из последних двух строк таблицы видно, что расхождение приближенного решения y и точного $\tilde{y}$ довольно значительно. Вообще с помощью конечно-разностных методов трудно получить решение с большой степенью точности.

Таблица 64

Решение краевой задачи (18) — (19) методом прогонки

i	0	1	2	3	4
c_i	—0,909	—0,899	—0,889	—0,878	—0,868
d_i	0	—0,004	—0,012	—0,023	—0,039
y_i	1,050	1,154	1,280	1,428	1,603
$\tilde{y}_i$	1	1,110	1,241	1,394	1,574

Продолжение таблицы 64

i	5	6	7	8	9	10
c_i	—0,856	—0,845	—0,833	—0,822	—0,810	
d_i	—0,058	—0,081	—0,109	—0,142	—0,180	
y_i	1,808	2,054	2,350	2,712	3,157	3,718
$\tilde{y}_i$	1,784	2,033	2,332	2,696	3,148	3,718

При требовании повышенной точности приходится применять более громоздкие приемы или вводить соответствующие поправки [1]. В частности, этот метод дает более точные результаты, если при переходе от краевых условий (2) к конечно-разностным соотношениям воспользоваться трехчленными формулами для производных в точках $x=a$ и $x=b$ (см. § 4):

$$y_0' = \frac{-y_2+4y_1-3y_0}{2h}, \quad y_n' = \frac{3y_n-4y_{n-1}+y_{n-2}}{2h}.$$

Тогда из краевых условий (2) будем иметь

$$\alpha_0 y_0 + \alpha_1 \frac{-y_2+4y_1-3y_0}{2h} = A, \quad \beta_0 y_0 + \beta_1 \frac{3y_n-4y_{n-1}+y_{n-2}}{2h} = B. \quad (20)$$

Для вычисления c_0 и d_0 берем первое краевое условие (20) и уравнение

$$y_2 + m_1 y_1 + n_1 y_0 = \hat{f}_1 h^2,$$

взятое из системы (3) при $i = 1$.

Исключая y_2 из этих двух уравнений, находим

$$y_0 = \frac{Rh - (m_1 + 4)\alpha_1 y_1}{(n_1 - 3)\alpha_1 + 2\alpha_0 h}, \tag{21}$$

где $R = 2A + \alpha_1 \hat{f}_1 h$. С другой стороны, из формулы (7) при $i = 0$ имеем

$$y_0 = c_0 (d_0 - y_1).$$

Сравнивая последнее равенство с (21), получаем

$$c_0 = \frac{(m_1 + 4)\alpha_1}{(n_1 - 3)\alpha_1 + 2\alpha_0 h}, \quad d_0 = \frac{Rh}{(m_1 + 4)\alpha_1}. \tag{22}$$

По известным значениям c_0 и d_0 последовательно определяются коэффициенты c_i, d_i $(i = 0, 1, \ldots, n-1)$ до c_{n-1} и d_{n-1} включительно (см. формулы (9)).

Обратный ход начинается с определения y_n.

Используя второе краевое условие (20) и формулы (7), взятые при $i = n-2$ и $i = n-1$, получим систему трех уравнений

$$\left.\begin{aligned} &\beta_0 y_n + \beta_1 \frac{3y_n - 4y_{n-1} + y_{n-2}}{2h} = B, \\ &y_{n-2} = c_{n-2}(d_{n-2} - y_{n-1}), \\ &y_{n-1} = c_{n-1}(d_{n-1} - y_n). \end{aligned}\right\} \tag{23}$$

Решая эту систему относительно y_n, будем иметь

$$y_n = \frac{2Bh + \beta_1 [(4 + c_{n-2}) c_{n-1} d_{n-1} - c_{n-2} d_{n-2}]}{\beta_1 [3 + 4c_{n-1} + c_{n-2} c_{n-1}] + 2\beta_0 h}. \tag{24}$$

Теперь по формулам (7) последовательно находим $y_{n-1}, y_{n-2}, \ldots, y_0$. Вычисления и в этом случае располагаем в виде таблицы 61.

Пример 3. Используя трехчленные формулы для производных в граничных точках, методом прогонки найти приближенное решение краевой задачи (18), (19).

Решение. От уравнений (18), (19) перейдем к конечно-разностным соотношениям

$$y_{i+1} + m_i y_i + n_i y_{i-1} = \hat{f}_i h^2, \qquad y_0 + \frac{y_2 - 4y_1 + 3y_2}{2h} = A, \qquad y_n = B, \tag{25}$$

где

$$h = 0{,}1; \qquad A = 0; \qquad B = 3{,}718,$$

$$m_i = -\frac{2 + 2h^2}{1 - ih^2}; \qquad n_i = \frac{1 + ih^2}{1 - ih^2}; \qquad \hat{f}_i h^2 = \varphi_i = -\frac{4ih}{1 - ih^2}.$$

Значения m_i, n_i и φ_i для $i = 0, 1, 2, 3, \ldots, 9$ приведены в таблице 65.

Таблица 65

Значения коэффициентов уравнения (25)

i	0	1	2	3	4	5	6	7	8	9
m_i	—2,020	—2,040	—2,060	—2,083	—2,105	—2,127	—2,149	—2,172	—2,196	—2,220
n_i	1	1,020	1,040	1,062	1,084	1,106	1,128	1,150	1,174	1,198
φ_i	0	—0,004	—0,008	—0,012	—0,017	—0,021	—0,025	—0,030	—0,035	—0,040

По формулам (22) находим значения c_0 и d_0, а затем по рекуррентным соотношениям (9) определяем c_i и d_i ($i=1, 2, \ldots, 9$). Эти значения приведены в первых двух строках таблицы 66. Обратный ход проводится по формулам (7) с учетом, что $y_{10}=3{,}718$.

Таблица 66

Решение краевой задачи (18)—(19) методом прогонки с использованием трехчленных формул для производных в концевых точках

i	0	1	2	3	4	5
c_i	— 0,899	— 0,890	— 0,883	— 0,873	— 0,863	— 0,853
d_i	0	— 0,005	— 0,013	— 0,024	— 0,040	— 0,059
y_i	1,003	1,114	1,247	1,399	1,578	1,789
$\tilde{y}_i$	1	1,110	1,241	1,394	1,574	1,784

Продолжение таблицы 66

i	6	7	8	9	10
c_i	— 0,842	— 0,831	— 0,820	— 0,809	
d_i	— 0,083	— 0,110	— 0,142	— 0,180	
y_i	2,038	2,337	2,702	3,153	3,718
$\tilde{y}_i$	2,033	2,332	2,696	3,148	3,718

Найденные значения $y_9, y_8, \ldots, y_0$ приведены в третьей строке таблицы 66. В последней строке указаны значения точного решения $\tilde{y}=x+e^{x^2}$.

Сравнивая таблицы 64 и 66, убеждаемся, что решение, полученное с помощью трехчленных формул (20), ближе к истинному, нежели в примере 2.

§ 6. Метод коллокации

Изложенный выше метод конечных разностей для решения краевых задач носит численный характер и позволяет получить таблицу значений искомой функции. Мы сейчас ознакомимся с методом, дающим возможность найти приближенное значение краевой задачи в виде аналитического выражения.

Пусть требуется определить функцию $y=y(x)$, удовлетворяющую линейному дифференциальному уравнению

$$L[y]\equiv y''+p(x)y'+q(x)y=f(x) \tag{1}$$

и линейным краевым условиям

$$\Gamma_a[y]\equiv\alpha_0 y(a)+\alpha_1 y'(a)=A, \qquad \Gamma_b[y]\equiv\beta_0 y(b)+\beta_1 y'(b)=B, \tag{2}$$

где $|\alpha_0|+|\alpha_1|\neq 0$, $|\beta_0|+|\beta_1|\neq 0$. Выберем некоторую совокупность линейно независимых функций

$$u_0(x),\ u_1(x),\ \ldots,\ u_n(x) \tag{3}$$

(базисные функции), из которых функция $u_0(x)$ удовлетворяет неоднородным краевым условиям

$$\Gamma_a[u_0]=A, \quad \Gamma_b[u_0]=B, \tag{4}$$

а остальные функции $u_i(x)$ $(i=1, 2, \ldots, n)$ удовлетворяют соответствующим однородным краевым условиям

$$\Gamma_a[u_i]=0, \quad \Gamma_b[u_i]=0 \quad (i=1, 2, \ldots, n). \tag{5}$$

Если краевые условия (2) однородны $(A=B=0)$, то можно положить $u_0(x)\equiv 0$ и рассматривать лишь систему функций

$$u_i(x) \quad (i=1, 2, \ldots, n).$$

Будем искать приближенное решение краевой задачи (1)—(2) в виде линейной комбинации базисных функций

$$y=u_0(x)+\sum_{i=1}^{n} C_i u_i(x). \tag{6}$$

Тогда функция y, очевидно, удовлетворяет краевым условиям (2). В самом деле, в силу линейности краевых условий имеем

$$\Gamma_a[y]=\Gamma_a[u_0]+\sum_{i=1}^{n} C_i\Gamma_a[u_i]=A+\sum_{i=1}^{n} C_i\cdot 0=A$$

и аналогично $\Gamma_b[y]=B$. Подставляя выражение (6) в уравнение (1), получим

$$R(x, C_1, C_2, \ldots, C_n) \equiv L[y]-f(x)=L[u_0]-f(x)+\sum_{i=1}^{n} C_i L[u_i]. \quad (7)$$

Если при некотором выборе коэффициентов C_i $(i=1, 2, \ldots, n)$ выполнено равенство

$$R(x, C_1, C_2, \ldots, C_n) \equiv 0 \text{ при } a \leqslant x \leqslant b,$$

то функция y является точным решением краевой задачи (1)—(2). Однако подобрать так удачно коэффициенты C_i, вообще говоря, невозможно. Поэтому ограничиваются тем, что требуют, чтобы функция $R(x, C_1, C_2, \ldots, C_n)$ обращалась в нуль в заданной достаточно густой системе точек $x_1, x_2, \ldots, x_n$ из $[a, b]$ (*точки коллокации*), в которых, таким образом, дифференциальное уравнение (1) будет удовлетворено точно; в качестве точек коллокации, например, можно выбрать точки, делящие отрезок $[a, b]$ на равные части. В результате получаем систему линейных уравнений

$$\left.\begin{array}{l} R(x_1, C_1, C_2, \ldots, C_n)=0, \\ \cdots\cdots\cdots\cdots \\ R(x_n, C_1, C_2, \ldots, C_n)=0. \end{array}\right\} \quad (8)$$

Из системы (8) в случае ее совместности обычным способом можно определить коэффициенты $C_1, C_2, \ldots, C_n$, после чего приближенное решение нашей краевой задачи дается формулой (6).

Пример 1. Методом коллокации решить краевую задачу [1]

$$y''+(1+x^2)y+1=0, \quad y(\pm 1)=0 \quad (9)$$

(см. § 4, пример 1).

Решение. В качестве базисных функций выберем полиномы $u_n(x)=x^{2n-2}(1-x^2)$ $(n=1, 2, \ldots)$, очевидно, удовлетворяющие краевым условиям $u_n(\pm 1)=0$. За точки коллокации возьмем $x_0=0$, $x_\pm=\pm 1/2$. Ограничиваясь двумя базисными функциями, положим

$$y=C_1(1-x^2)+C_2(x^2-x^4).$$

Подстановка в дифференциальное уравнение (9) дает

$$R(x)=-2C_1+C_2(2-12x^2)+(1+x^2)[C_1(1-x^2)+ \\ +C_2(x^2-x^4)]+1=1-C_1(1+x^4)+C_2(2-11x^2-x^6). \quad (10)$$

В точках коллокации $x_0=0$ и $x_\pm=\pm 1/2$ имеем $R(x_0)=0$, $R(x_\pm)=0$. Отсюда, используя формулу (10), получаем для определения коэффициентов C_1 и C_2 линейную систему уравнений

$$1-C_1+2C_2=0, \qquad 1-\frac{17}{16}C_1-\frac{49}{64}C_2=0. \quad (11)$$

Решив систему (11), находим $C_1=0{,}957$; $C_2=-0{,}022$. Следовательно,

имеем приближенное решение

$$y \approx 0{,}957\,(1-x^2)-0{,}022\,(x^2-x^4)=0{,}957-0{,}979x^2+0{,}022x^4.$$

В частности, получаем $y(0)=0{,}957$. Для сравнения приводим соответствующее значение $y(0)$, полученное методом конечных разностей (см. § 4): $y(0)=y_0=0{,}967$.

§ 7. Метод наименьших квадратов

Снова рассмотрим краевую задачу

$$L[y]=f(x), \qquad \Gamma_a[y]=A, \qquad \Gamma_b[y]=B, \tag{1}$$

где смысл сокращенных обозначений был установлен в предыдущем параграфе. Придерживаясь обозначений § 6, полагаем

$$y=u_0(x)+\sum_{i=1}^{n} C_i u_i(x), \tag{2}$$

где функции $u_i=u_i(x)$ таковы, что

$$\Gamma_a[u_i]=\begin{cases} A \text{ при } i=0, \\ 0 \text{ при } i>0, \end{cases} \qquad \Gamma_b[u_i]=\begin{cases} B \text{ при } i=0, \\ 0 \text{ при } i>0. \end{cases}$$

Подставляя выражение (2) в дифференциальное уравнение (1), получаем невязку

$$R(x, C_1, C_2, \ldots, C_n)=L[u_0]-f(x)+\sum_{i=1}^{n} C_i L[u_i], \tag{3}$$

которая должна быть при $a \leqslant x \leqslant b$ как можно меньше по абсолютной величине. Поэтому выдвигаем требование, чтобы

$$I=\int_a^b R^2(x, C_1, C_2, \ldots, C_n)\,dx=\min \tag{4}$$

(интегральный метод наименьших квадратов).

Для минимума интеграла I необходимо выполнение условий

$$\left.\begin{aligned} \frac{1}{2}\frac{\partial I}{\partial C_1}&=\int_a^b R\frac{\partial R}{\partial C_1}\,dx=0, \\ \frac{1}{2}\frac{\partial I}{\partial C_2}&=\int_a^b R\frac{\partial R}{\partial C_2}\,dx=0, \\ &\ldots\ldots\ldots\ldots \\ \frac{1}{2}\frac{\partial I}{\partial C_n}&=\int_a^b R\frac{\partial R}{\partial C_n}\,dx=0. \end{aligned}\right\} \tag{5}$$

В результате получается система линейных уравнений относительно коэффициентов $C_1, C_2, \ldots, C_n$, из которой и определяются эти последние.

Пример 1. Интегральным методом наименьших квадратов решить краевую задачу

$$y''+(1+x^2)\,y+1=0, \quad y(\pm 1)=0, \tag{6}$$

разобранную в § 6.

Решение. Полагая $u_1=1-x^2$, $u_2=x^2-x^4$, будем иметь

$$y=C_1(1-x^2)+C_2(x^2-x^4).$$

Подставляя это выражение в уравнение (6), получаем невязку (см. § 6, формула (10))

$$R(x)=1-(1+x^4)\,C_1+(2-11x^2-x^6)\,C_2. \tag{7}$$

Ввиду симметрии задачи в качестве основного отрезка достаточно рассматривать отрезок [0, 1].

Согласно приведенному выше способу наименьших квадратов составляем выражение

$$I=\int_0^1 R^2(x)\,dx=\int_0^1 [1-(1+x^4)\,C_1+(2-11x^2-x^6)\,C_2]^2\,dx$$

и подбираем коэффициенты C_1 и C_2 так, чтобы интеграл I имел наименьшее значение. Это дает нам систему уравнений

$$\left.\begin{aligned}
&\frac{1}{2}\frac{\partial I}{\partial C_1}=-\int_0^1 (1+x^4)\,[1-(1+x^4)\,C_1+(2-11x^2-x^6)\,C_2]\,dx=0,\\
&\frac{1}{2}\frac{\partial I}{\partial C_2}=\int_0^1 (2-11x^2-x^6)\,[1-(1+x^4)\,C_1+\\
&\qquad\qquad +(2-11x^2-x^6)\,C_2]\,dx=0
\end{aligned}\right\}$$

или

$$\frac{68}{45}C_1+\frac{3548}{1155}C_2=\frac{5}{4}, \qquad \frac{3548}{1155}C_1+\frac{63404}{4095}C_2=\frac{38}{21}.$$

Отсюда $C_1=0{,}985$; $C_2=-0{,}078$. Следовательно,

$$y=0{,}985\,(1-x^2)-0{,}078\,(x^2-x^4). \tag{8}$$

В частности, $y(0)=0{,}985$ (ср. § 6).

Вместо минимума интеграла (7) можно искать минимум конечной суммы *(точечный метод наименьших квадратов)*

$$I_N=\sum_{i=1}^{N} R^2(x_i,\ C_1,\ C_2,\ \ldots,\ C_n),$$

где $x_1, x_2, \ldots, x_N$ — некоторая достаточно густая система точек отрезка $[a, b]$ (для простоты обычно выбираются равноотстоящие точки).

При применении точечного метода наименьших квадратов следует полагать $N>n$, т. е. число точек x_i $(i=1, 2, \ldots, N)$ должно

быть больше числа параметров C_j ($j=1, 2, \ldots, n$). В случае, если $N=n$, параметры C_j можно определить из системы уравнений

$$R(x_i, C_1, C_2, \ldots, C_n)=0 \quad (i=1, 2, \ldots, n),$$

т. е. мы приходим к методу коллокации (§ 6).

Пример 2. Точечным методом наименьших квадратов найти решение краевой задачи (6), выбирая узлы

$$x_0=0, \quad x_{\pm 1}=\pm\frac{1}{4}, \quad x_{\pm 2}=\pm\frac{1}{2}, \quad x_{\pm 3}=\pm\frac{3}{4}.$$

Решение. Ввиду симметрии задачи за основной промежуток примем отрезок [0, 1]. Положим

$$y=C_1(1-x^2)+C_2(x^2-x^4);$$

тогда невязка $R(x)$ выражается формулой (7). Пусть

$$R(x_i)=1+\alpha_i C_1+\beta_i C_2 \quad (i=0, 1, 2, 3),$$

где $\alpha_i=-(1+x_i^4)$, $\beta_i=2-11x_i^2-x_i^6$; тогда

$$I=\sum_{i=0}^{3} R^2(x_i)=\sum_{i=0}^{3}(1+\alpha_i C_1+\beta_i C_2)^2.$$

Используя необходимые условия для минимума функции I, получим нормальную систему

$$\left.\begin{aligned} \frac{1}{2}\frac{\partial I}{\partial C_1}&=\sum_i \alpha_i+C_1\sum_i \alpha_i^2+C_2\sum_i \alpha_i\beta_i=0,\\ \frac{1}{2}\frac{\partial I}{\partial C_2}&=\sum_i \beta_i+C_1\sum_i \alpha_i\beta_i+C_2\sum_i \beta_i^2=0.\end{aligned}\right\} \tag{9}$$

Соответствующие числовые данные, вычисленные с точностью до 10^{-4}, приведены в таблице 67.

Таблица 67

Коэффициенты нормальной системы (9)

x	α	β	α^2	$\alpha\beta$	β^2
0	−1	2	1	−2	4
$\frac{1}{4}$	−1,0039	1,3123	1,0078	−1,3174	1,7222
$\frac{1}{2}$	−1,0625	−0,7656	1,1289	0,8134	0,5861
$\frac{3}{4}$	−1,3164	−4,3655	1,7329	5,7465	19,0576
$\sum$	−4,3828	−1,8188	4,8696	3,2425	25,3659

Отсюда для определения коэффициентов C_1 и C_2 получаем систему

$$\left.\begin{aligned} 4{,}8696C_1 + 3{,}2425C_2 &= 4{,}3828, \\ 3{,}2425C_1 + 25{,}3659C_2 &= 1{,}8188. \end{aligned}\right\} \tag{10}$$

Решив систему (10), будем иметь $C_1 = 0{,}9317$; $C_2 = -0{,}0474$, и, следовательно,

$$y = 0{,}9317\,(1 - x^2) - 0{,}0474\,(x^2 - x^4).$$

§ 8. Метод Галеркина

Метод Галеркина основан на одной теореме из теории общих рядов Фурье.

Теорема. *Пусть $\{u_n(x)\}$ — полная система функций с ненулевой нормой, ортогональных на отрезке $[a, b]$. Если непрерывная функция $f(x)$ ортогональна на отрезке $[a, b]$ ко всем функциям $u_n(x)$, т. е.*

$$\int_a^b f(x)\,u_n(x)\,dx = 0 \quad (n = 0,\ 1,\ 2,\ \ldots), \tag{1}$$

то $f(x) \equiv 0$ при $a \leqslant x \leqslant b$.

Доказательство. Рассмотрим ряд Фурье функции $f(x)$ относительно заданной системы ортогональных функций

$$f(x) \sim \sum_{n=0}^{\infty} c_n u_n(x). \tag{2}$$

Как известно [5], коэффициенты Фурье c_n определяются по формуле

$$c_n = \frac{1}{\|u_n\|^2} \int_a^b f(x)\,u_n(x)\,dx, \quad \text{где} \quad \|u_n\|^2 = \int_a^b u_n^2(x)\,dx > 0.$$

В силу условия (1) имеем

$$c_n = 0 \quad (n = 0,\ 1,\ 2,\ \ldots). \tag{3}$$

Для полной системы $\{u_n(x)\}$ по отношению к любой непрерывной функции $f(x)$ выполнено равенство полноты [4]

$$\int_a^b f^2(x)\,dx = \sum_{n=0}^{\infty} \|u_n\|^2 c_n^2. \tag{4}$$

Отсюда, учитывая равенство (3), имеем

$$\int_a^b f^2(x)\,dx = 0,$$

и, следовательно, $f(x) \equiv 0$ при $a \leqslant x \leqslant b$.

Замечание. Из формулы (4) вытекает, что если непрерывная функция $f(x)$ ортогональна к конечной системе функций $u_0(x)$, $u_1(x)$, $u_2(x)$, ..., $u_N(x)$ (т. е. $c_0=c_1=c_2=\ldots=c_N=0$), то

$$\int_a^b f^2(x)\,dx=\sum_{n=N+1}^{\infty}\|u_n\|^2 c_n^2<\varepsilon$$

при N достаточно большом. В этом случае функция $f(x)$ *в среднем* на отрезке $[a, b]$ будет сколь угодно мала. При дополнительных ограничениях отсюда следует, что $|f(x)|$ также мал на отрезке $a\leqslant x\leqslant b$.

Перейдем теперь к изложению метода Галеркина. Пусть имеем линейную краевую задачу (см. § 2)

$$L[y]=f(x), \tag{5}$$

где $L[y]=y''+p(x)y'+q(x)y$, при наличии линейных краевых условий

$$\Gamma_a[y]\equiv\alpha_0 y(a)+\alpha_1 y'(a)=A;\qquad \Gamma_b[y]\equiv\beta_0 y(b)+\beta_1 y'(b)=B \tag{6}$$
$$(|\alpha_0|+|\alpha_1|\neq 0,\quad |\beta_0|+|\beta_1|\neq 0).$$

Выберем конечную систему базисных функций $\{u_i(x)\}$ $(i=0, 1, \ldots, n)$, составляющих часть некоторой полной системы, причем позаботимся, чтобы функция $u_0(x)$ удовлетворяла неоднородным краевым условиям

$$\Gamma_a[u_0]=A,\quad \Gamma_b[u_0]=B,$$

а функции $u_i(x)$ $(i=1, 2, \ldots, n)$ удовлетворяли бы однородным краевым условиям

$$\Gamma_a[u_i]=\Gamma_b[u_i]=0\qquad (i=1, 2, \ldots, n).$$

Решение краевой задачи (5)—(6) будем, как обычно, искать в виде

$$y=u_0(x)+\sum_{i=1}^{n}C_i u_i(x). \tag{7}$$

При нашем подборе базисных функций $u_i(x)$ функция y, определяемая формулой (7), очевидно, удовлетворяет краевым условиям (6) при любом выборе коэффициентов C_i. Выражение (7) подставим в дифференциальное уравнение (5), что дает *невязку*

$$R(x, C_1, C_2, \ldots, C_n)=L[u_0]+\sum_{i=1}^{n}C_i L[u_i]-f(x).$$

Для точного решения y нашей краевой задачи функция $R\equiv 0$; поэтому для получения приближенного решения, близкого к точному, нам выгодно подобрать коэффициенты C_i так, чтобы функция R была в каком-то смысле мала.

Согласно методу Галеркина требуем, чтобы невязка R была ортогональна к базисным функциям $u_i(x)$ $(i=1, 2, \ldots, n)$, что при достаточно большом числе этих функций, в силу приведенного выше замечания, обеспечивает малость невязки в среднем.

Насколько это приближенное решение близко к точному, в общем случае вопрос остается открытым. Таким образом, для определения коэффициентов C_i $(i=1, 2, \ldots, n)$ приходим к системе линейных уравнений

$$\left.\begin{aligned}
&\int_a^b u_1(x)\,R(x, C_1, C_2, \ldots, C_n)\,dx=0,\\
&\int_a^b u_2(x)\,R(x, C_1, C_2, \ldots, C_n)\,dx=0,\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
&\int_a^b u_n(x)\,R(x, C_1, C_2, \ldots, C_n)\,dx=0
\end{aligned}\right\}$$

или, более подробно,

$$\sum_{i=1}^{n} C_i \int_a^b u_i(x)\,L[u_i]\,dx=\int_a^b u_i(x)\,\{f(x)-L[u_0]\}\,dx$$
$$(i=1, 2, \ldots, n).$$

Достаточные условия сходимости метода Галеркина приведены в книге Михлина [4].

Пример 1. Методом Галеркина найти приближенное решение уравнения [2]

$$y''+xy'+y=2x, \tag{8}$$

удовлетворяющее краевым условиям

$$y(0)=1, \quad y(1)=0. \tag{9}$$

Решение. В качестве системы базисных функций $u_i(x)$ $(i=0, 1, 2, 3)$ выбираем следующие функции:

$$u_0(x)=1-x, \quad u_1(x)=x(1-x),$$
$$u_2(x)=x^2(1-x), \quad u_3(x)=x^3(1-x).$$

Приближенное решение задачи ищем в виде полинома

$$y=(1-x)+C_1x(1-x)+C_2x^2(1-x)+C_3x^3(1-x).$$

Подставляя y в левую часть заданного дифференциального уравнения (8), получаем невязку

$$R(x, C_1, C_2, C_3)=(1-4x)+C_1(-2+2x-3x^2)+$$
$$+C_2(2-6x+3x^2-4x^3)+C_3(6x-12x^2+4x^3-5x^4). \tag{10}$$

Условия ортогональности функции R к функциям $u_1(x)$, $u_2(x)$, $u_3(x)$ приводят к системе

$$\left.\begin{aligned}&\int_0^1 (x-x^2)\,R(x,\ C_1,\ C_2,\ C_3)\,dx=0,\\&\int_0^1 (x^2-x^3)\,R(x,\ C_1,\ C_2,\ C_3)\,dx=0,\\&\int_0^1 (x^3-x^4)\,R(x,\ C_1,\ C_2,\ C_3)\,dx=0.\end{aligned}\right\}$$

Подставляя вместо $R(x)$ его значение (10), после соответствующего интегрирования получаем систему

$$\left.\begin{aligned}133C_1+\ 63C_2+\ 36C_3&=-\ 70,\\140C_1+108C_2+\ 79C_3&=-\ 98,\\264C_1+252C_2+211C_3&=-210.\end{aligned}\right\}$$

Отсюда находим $C_1=-0{,}2090$; $C_2=-0{,}7894$; $C_3=0{,}2090$, и, следовательно,

$$y=(1-x)\,(1-0{,}2090x-0{,}7894x^2+0{,}2090x^3).$$

§ 9. Понятие о приближенных методах решения общей краевой задачи

Рассмотрим общее дифференциальное уравнение n-го порядка

$$F(x,\ y,\ y',\ \ldots,\ y^{(n)})=0 \tag{1}$$

с заданными, вообще говоря, нелинейными краевыми условиями

$$V_\nu(y_1,\ y_1',\ \ldots,\ y_1^{(\sigma_1)};\ \ldots;\ y_k,\ y_k',\ \ldots,\ y_k^{(\sigma_k)})=A_\nu \tag{2}$$
$$(\nu=1,\ 2,\ \ldots,\ n),$$

где $y_i^{(s)}=y^{(s)}(x_i)$ $(s=0,\ 1,\ 2,\ \ldots,\ \sigma_i)$ и система точек $x_1, x_2, \ldots, x_k$ задана (см. § 1).

Для приближенного решения краевой задачи (1)—(2) выбирают функцию

$$Y=Y(x,\ C_1,\ C_2,\ \ldots,\ C_p), \tag{3}$$

содержащую независимые параметры $C_1,\ C_2,\ \ldots,\ C_p$ и такую, что при любом выборе этих параметров функция Y удовлетворяет краевым условиям (2). Подставляя выражение Y в левую часть данного дифференциального уравнения (1), получаем *невязку*

$$R(x,\ C_1,\ C_2,\ \ldots,\ C_p)\equiv F(x,\ Y,\ Y',\ \ldots,\ Y^{(n)}). \tag{4}$$

Наша цель состоит в том, чтобы сделать функцию R наименее уклоняющейся от нуля в каком-то смысле для интересующей нас области.

Различная реализация этого требования приводит к различным методам решения краевой задачи.

Мы здесь рассмотрим наиболее употребительные из этих методов [1].

I. Метод коллокации. Выбираем систему точек $\xi_1, \xi_2, \ldots, \xi_p$ (*точки коллокации*), по возможности учитывая при этом особенности уравнения (1), и требуем, чтобы в этих точках дифференциальное уравнение (1) удовлетворялось точно. Это дает систему

$$\left.\begin{aligned} &R(\xi_1, C_1, C_2, \ldots, C_p)=0,\\ &R(\xi_2, C_1, C_2, \ldots, C_p)=0,\\ &\ldots\ldots\ldots\ldots\ldots\ldots\\ &R(\xi_p, C_1, C_2, \ldots, C_p)=0, \end{aligned}\right\} \tag{5}$$

из которой, вообще говоря, можно определить параметры $C_1, C_2, \ldots, C_p$ (ср. § 6).

II. Метод наименьших квадратов. Отыскиваются значения $C_1, C_2, \ldots, C_p$, дающие минимум интегралу (*интегральный метод наименьших квадратов*)

$$I=\int_a^b R^2(x, C_1, C_2, \ldots, C_p)\,dx,$$

где отрезок $[a, b]$ содержит заданные точки x_i $(i=1, 2, \ldots, k)$. Этод метод приводит к системе уравнений (ср. § 7)

$$\frac{\partial I}{\partial C_1}=0, \quad \frac{\partial I}{\partial C_2}=0, \quad \ldots, \quad \frac{\partial I}{\partial C_p}=0.$$

Вместо интеграла I можно брать конечную сумму (*точечный метод наименьших квадратов*)

$$I_N=\sum_{i=1}^{N} R^2(\xi_i, C_1, C_2, \ldots, C_p),$$

где ξ_i $(i=1, 2, \ldots, N;\ N\geqslant p)$ — выбранная система точек.

III. Метод Галеркина. Если краевые условия (2) линейные (дифференциальное уравнение (1) при этом может быть нелинейным), то можно положить

$$Y=u_0(x)+\sum_{i=1}^{n} C_i u_i(x),$$

причем функция Y должна удовлетворять краевым условиям (2) при любом выборе коэффициентов C_i $(i=1, 2, \ldots, n)$.

Согласно методу Галеркина коэффициенты C_i определяются из условия ортогональности невязки R к функциям $u_1(x), u_2(x), \ldots, u_n(x)$,

т. е. из системы линейных уравнений

$$\left.\begin{aligned}
&\int_a^b u_1(x)\,R(x, C_1, C_2, \ldots, C_n)\,dx = 0,\\
&\int_a^b u_2(x)\,R(x, C_1, C_2, \ldots, C_n)\,dx = 0,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\\
&\int_a^b u_n(x)\,R(x, C_1, C_2, \ldots, C_n)\,dx = 0.
\end{aligned}\right\} \qquad (6)$$

IV. Обобщенный метод Галеркина (*метод ортогональных проекций*). Выбираем p линейно независимых функций $\varphi_1(x), \varphi_2(x), \ldots, \varphi_p(x)$ (по числу параметров) и определяем параметры $C_1, C_2, \ldots, C_p$ из условия ортогональности функции R ко всем этим функциям, т. е. постоянные C_i определяются из системы

$$\left.\begin{aligned}
&\int_a^b \varphi_1(x)\,R(x, C_1, C_2, \ldots, C_p)\,dx = 0,\\
&\int_a^b \varphi_2(x)\,R(x, C_1, C_2, \ldots, C_p)\,dx = 0,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\\
&\int_a^b \varphi_p(x)\,R(x, C_1, C_2, \ldots, C_p)\,dx = 0.
\end{aligned}\right\} \qquad (7)$$

В качестве функций $\varphi_i(x)$ обычно берут первые p функций из какой-нибудь полной системы, например системы степеней x или системы тригонометрических полиномов и т. п.

V. Метод подобластей. Разбиваем отрезок $[a, b]$, содержащий данные точки x_i $(i = 1, 2, \ldots, k)$, на p частичных отрезков $[\xi_0, \xi_1]$, $[\xi_1, \xi_2], \ldots, [\xi_{p-1}, \xi_p]$, где $\xi_0 = a$ и $\xi_p = b$. Постоянные C_i $(i = 1, 2, \ldots, p)$ определяются из системы уравнений

$$\int_{\xi_{j-1}}^{\xi_j} R(x, C_1, C_2, \ldots, C_p)\,dx = 0 \qquad (j = 1, 2, \ldots, p).$$

Кроме того, применяют также *вариационные методы* решения краевых задач (см. гл. VI). Сопоставление этих методов см. [1].

В заключение отметим, что имеются и другие методы решения краевых задач для обыкновенных дифференциальных уравнений. Некоторые из них: метод возмущений, метод разностной факторизации и

другие — хорошо описаны в книге [6]. Там же рассмотрен вопрос и об оценке погрешности конечно-разностных методов при решении краевых задач. Можно также порекомендовать переведенную на русский язык книгу Дж. Н. Ланса [8].

ЛИТЕРАТУРА К ГЛАВЕ IV

[1] Коллатц Л., Численные методы решения дифференциальных уравнений, ИЛ, 1953, гл. II.
[2] Милн В. Э., Численное решение дифференциальных уравнений, ИЛ, 1955, гл. VII.
[3] Эльсгольц Л. Э., Вариационное исчисление, Гостехиздат, 1954.
[4] Михлин С. Г., Вариационные методы в математической физике, Гостехиздат, 1957, гл. III, V, IX.
[5] Толстов Г. П., Ряды Фурье, изд. 2, Физматгиз, 1960, гл II.
[6] Положий Г. Н. и др., Математический практикум, Физматгиз, 1960.
[7] Локуциевский О. В., Успехи матем. наук, 11, вып. 3 (69) (1956).
[8] Ланс Дж. Н., Численные методы для быстродействующих вычислительных машин, ИЛ, 1962.

ГЛАВА V

ПРИБЛИЖЕННЫЕ МЕТОДЫ РЕШЕНИЯ КРАЕВЫХ ЗАДАЧ ДЛЯ ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ С ЧАСТНЫМИ ПРОИЗВОДНЫМИ

§ 1. Классификация дифференциальных уравнений с частными производными

В этой главе будут рассмотрены приближенные методы решения некоторых задач для дифференциальных уравнений с частными производными второго порядка с двумя независимыми переменными. В общем случае такое уравнение имеет вид

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

где x, y — независимые переменные, u — искомая функция, u_x, u_y, u_{xx}, u_{xy}, u_{yy} — ее первые и вторые частные производные по аргументам x и y (для удобства записи производных «штрихи» опускаются).

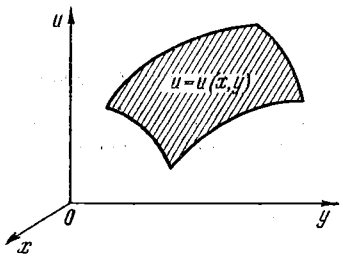

Рис. 49.

Решением уравнения (1) называется функция $u=u(x,\ y)$, обращающая это уравнение в тождество. График решения представляет собой поверхность в пространстве $Oxyu$ (*интегральная поверхность*) (рис. 49).

Пример 1. Решить уравнение

$$\frac{\partial^2 u}{\partial y^2}=0.$$

Интегрируя это уравнение по y два раза, будем иметь $u=y\varphi(x)+\psi(x)$, где $\varphi(x)$ и $\psi(x)$ — произвольные функции. Интегральные поверхности представляют собой линейчатые поверхности, образующие которых параллельны координатной плоскости Oyu.

Уравнение (1) называется *линейным*, точнее, *вполне линейным*, если оно первой степени относительно искомой функции и всех ее

производных и не содержит их произведений, т. е. если это уравнение может быть записано в виде

$$A\frac{\partial^2 u}{\partial x^2}+2B\frac{\partial^2 u}{\partial x\,\partial y}+C\frac{\partial^2 u}{\partial y^2}+a\frac{\partial u}{\partial x}+b\frac{\partial u}{\partial y}+cu=F(x,\ y), \tag{2}$$

причем коэффициенты A, B, C, a, b, c могут зависеть лишь от x и y. В частности, если эти коэффициенты не зависят от x и y, то уравнение (2) представляет собой *линейное дифференциальное уравнение с постоянными коэффициентами*. Остановимся подробнее на случае линейного уравнения (2).

Пусть $D=AC-B^2$ — *дискриминант уравнения*. В зависимости от знака функции D линейное дифференциальное уравнение (2) относится в данной области к одному из следующих типов:

$D>0$ — *эллиптический тип*;

$D=0$ — *параболический тип*;

$D<0$ — *гиперболический тип*;

D не сохраняет постоянного знака — *смешанный тип*.

Тип линейного уравнения (2) является его важной особенностью и сохраняется при любом невырожденном преобразовании

$$\xi=\varphi(x,\ y),\quad \eta=\psi(x,\ y), \tag{3}$$

т. е. таком, что якобиан

$$\frac{\partial(\varphi,\ \psi)}{\partial(x,\ y)}\neq 0.$$

Пример 2. Температура $u=u(x,\ y)$ точки $(x,\ y)$ пластинки при стационарном распределении (т. е. при распределении, не зависящем от времени) и отсутствии источников тепла удовлетворяет уравнению Лапласа [1]—[5]

$$\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}=0. \tag{4}$$

Здесь $A=1$, $B=0$, $C=1$ и $D=AC-B^2>0$, т. е. уравнение (4) эллиптического типа.

Пример 3. Температура $u=u(x,\ t)$ точки однородного тонкого стержня с абсциссой x для каждого момента времени t удовлетворяет одномерному уравнению теплопроводности [4], [5]

$$\frac{\partial u}{\partial t}-a^2\frac{\partial^2 u}{\partial x^2}=F(x,\ t), \tag{5}$$

где a — постоянная, зависящая от физических свойств стержня, и $F(x,\ t)$ — функция, связанная с плотностью источников распределения тепла. Если в стержне отсутствуют источники тепла, то уравнение теплопроводности имеет вид

$$\frac{\partial u}{\partial t}=a^2\frac{\partial^2 u}{\partial x^2}.$$

Введя новое время $a^2t=\tau$, получаем *приведенное уравнение теплопроводности*

$$\frac{\partial u}{\partial \tau}=\frac{\partial^2 u}{\partial x^2}. \tag{6}$$

Уравнения теплопроводности (5) и (6), очевидно, параболического типа.

Пример 4. Поперечное смещение $u=u(x, t)$ точки однородной струны с абсциссой x (рис. 50) в случае наличия внешней силы для каждого момента t удовлетворяет неоднородному одномерному *волновому уравнению* [1]—[5]

$$\frac{\partial^2 u}{\partial t^2}-a^2\frac{\partial^2 u}{\partial x^2}=F(x, t), \tag{7}$$

где a—постоянная и $F(x, t)$—функция, зависящая от внешней силы. Уравнение (7) носит название *уравнения колебаний струны*. Если внешняя сила отсутствует (свободные колебания), то уравнение колебаний струны имеет вид

$$\frac{\partial^2 u}{\partial t^2}=a^2\frac{\partial^2 u}{\partial x^2}. \tag{8}$$

Рис. 50.

Уравнения колебаний струны (7) и (8) относятся к гиперболическому типу.

С линейным дифференциальным уравнением

$$A\frac{\partial^2 u}{\partial x^2}+2B\frac{\partial^2 u}{\partial x\,\partial y}+C\frac{\partial^2 u}{\partial y^2}+a\frac{\partial u}{\partial x}+b\frac{\partial u}{\partial y}+cu=F(x, y) \tag{9}$$

связано обыкновенное дифференциальное уравнение

$$A(dy)^2-2B\,dx\,dy+C(dx)^2=0, \tag{10}$$

называемое *характеристическим*; решения уравнения (10) называются *характеристиками уравнения* (9).

Для уравнения (9) гиперболического типа существует два семейства характеристик

$$\varphi(x, y)=C_1$$

и

$$\psi(x, y)=C_2$$

(рис. 51). Производя в уравнении (9) преобразование $\xi=\varphi(x, y)$, $\eta=\psi(x, y)$, т. е. принимая параметры этих семейств за новые криволинейные координаты, будем иметь канонический вид уравнения гиперболического типа:

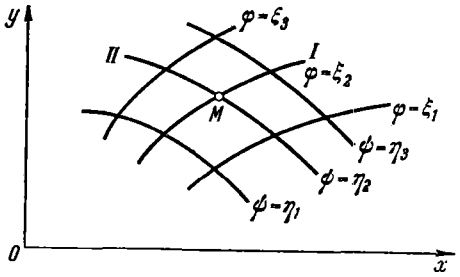

Рис. 51.

$$u_{\xi\eta}+\alpha(\xi, \eta)\,u_\xi+\beta(\xi, \eta)\,u_\eta+\gamma(\xi, \eta)\,u=f(\xi, \eta).$$

Уравнение (9) параболического типа имеет одно семейство характеристик

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

В результате преобразования $\xi=\varphi(x, y)$, $\eta=y$ уравнение парабо-

лического типа приводится к каноническому виду:

$$u_{\eta\eta}+\alpha(\xi,\ \eta)\,u_{\xi}+\beta(\xi,\ \eta)\,u_{\eta}+\gamma(\xi,\ \eta)\,u=f(\xi,\ \eta).$$

Наконец, уравнение (9) эллиптического типа допускает два семейства комплексных характеристик:

$$\varphi(x,\ y)+i\psi(x,\ y)=C_1,\quad \varphi(x,\ y)-i\psi(x,\ y)=C_2.$$

Производя преобразования $\xi=\varphi(x,\ y)$, $\eta=\psi(x,\ y)$, получим канонический вид уравнения эллиптического типа:

$$\Delta u+\alpha(\xi,\ \eta)\,u_{\xi}+\beta(\xi,\ \eta)\,u_{\eta}+\gamma(\xi,\ \eta)\,u=f(\xi,\ \eta),$$

где $\Delta u=u_{\xi\xi}+u_{\eta\eta}$ — *оператор Лапласа*.

Простейшее уравнение эллиптического типа $\Delta u=0$ носит название *уравнения Лапласа*. Неоднородное уравнение Лапласа $\Delta u=f(\xi,\ \eta)$ называется *уравнением Пуассона*.

§ 2. Начальные и краевые условия. Задача Коши. Смешанная задача. Корректность постановки смешанной задачи

Дифференциальное уравнение с частными производными имеет в общем случае бесчисленное множество решений. Поэтому, если физический процесс описывается с помощью уравнения с частными производными, то для однозначной характеристики этого процесса нужно к уравнению присоединить какие-то дополнительные условия. Эти дополнительные данные в простейшем случае состоят из *начальных* и *краевых* (граничных) условий. В сущности, различить эти условия можно лишь в том случае, если одна из независимых переменных дифференциального уравнения играет роль времени, а другая — пространственной координаты (для случая двух независимых переменных). При этом условия, относящиеся к начальному моменту времени, называются *начальными*, а условия, относящиеся к фиксированным значениям координат (обычно это координаты граничных точек рассматриваемого линейного континуума) — *краевыми*.

Рис. 52.

Пример 1. Пусть имеется теплоизолированный (кроме, может быть, концов) однородный нагретый стержень $0\leqslant x\leqslant l$, где l — длина стержня (рис. 52). Температура стержня $u=u(x,\ t)$ в точке $x\ (0<x<l)$ для любого момента времени t удовлетворяет уравнению теплопроводности (см. § 1, пример 3)

$$\frac{\partial u}{\partial t}=a^2\frac{\partial^2 u}{\partial x^2},\tag{1}$$

где a — постоянная.

В начальный момент $t=t_0$ для внутренних точек стержня обычно задается начальное распределение температуры. Это приводит к начальному условию

$$u(x, t_0)=f(x) \tag{2}$$

при $0<x<l$, где $f(x)$ — известная функция. Условие (2) не обеспечивает однозначности решения дифференциального уравнения (1), так как физически ясно, что распределение температуры $u(x, t)$ в стержне для последующих моментов времени $t>t_0$ существенно зависит от того, в каком состоянии находятся концы стержня $x=0$ и $x=l$ (есть ли там утечка тепла, каков тепловой режим и т. п.).

В зависимости от состояния конца $x=0$ имеем следующие основные краевые условия.

1. *Конец стержня $x=0$ поддерживается при заданной температуре*

$$u(0, t)=\varphi(t), \tag{3}$$

где $\varphi(t)$ — известная функция. В частности, если эта температура равна нулю, то краевое условие имеет вид

$$u(0, t)=0. \tag{4}$$

2. *Конец стержня $x=0$ теплоизолирован*, т. е. утечка тепла в окружающую среду отсутствует:

$$u_x(0, t)=0. \tag{5}$$

3. *На конце стержня $x=0$ происходит лучеиспускание тепла в окружающую среду*, температура которой меняется по заданному закону

$$u(0, t)+\alpha u_x(0, t)=\varphi(t), \tag{6}$$

где α — постоянная и $\varphi(t)$ — известная функция. В частности, если температура внешней среды равна нулю, то получим

$$u(0, t)+\alpha u_x(0, t)=0. \tag{7}$$

Смешанное краевое условие (6) в некотором смысле можно считать общим, а именно, полагая $\alpha=0$, получим краевое условие (3), а при $\alpha=\infty$ будем иметь краевое условие (5). Возможны и другие типы краевых условий. Аналогичные краевые условия могут быть также для конца $x=l$. Комбинируя краевые условия для концов $x=0$ и $x=l$, будем иметь краевые задачи для стержня, которые при наличии начального условия (2), вообще говоря, имеют единственные решения.

Пример 2. Рассмотрим свободные колебания однородной ограниченной струны длины l $(0<x<l)$. Поперечное смещение $u=u(x, t)$ при $0<x<l$ для любого момента времени t удовлетворяет волновому уравнению (см. § 1, пример 4)

$$\frac{\partial^2 u}{\partial t^2}=a^2\frac{\partial^2 u}{\partial x^2}, \tag{8}$$

где a — постоянная. В начальный момент $t=t_0$ обычно задаются форма струны и распределение скоростей ее точек. Это дает начальные условия

$$u(x, t_0)=\varphi(x), \quad u_t(x, t_0)=\varphi_1(x), \tag{9}$$

где $\varphi(x)$ и $\varphi_1(x)$ — известные функции, определенные в интервале $0<x<l$. В зависимости от способов заделки концов струны $x=0$ и $x=l$ будем иметь следующие основные краевые условия.

1. *Конец жестко закреплен:*

$$u(0, t)=0 \text{ (или } u(l, t)=0). \tag{10'}$$

2. *Конец упруго закреплен:*

$$u_x(0, t)-k_1u(0, t)=0 \text{ (или } u_x(l, t)+k_2u(l, t)=0), \tag{10''}$$

где k_1 и k_2 — положительные постоянные.

3. *Конец свободен:*

$$u_x(0, t)=0 \text{ (или } u_x(l, t)=0). \tag{10'''}$$

При достаточной гладкости функций $\varphi(x)$ и $\varphi_1(x)$ задача (9)—(10) имеет единственное решение.

Рассмотрим общую постановку задачи с начальными условиями. Пусть дано линейное дифференциальное уравнение

$$L[u]=F(x, y), \tag{11}$$

где

$$L[u]=A(x, y)\frac{\partial^2 u}{\partial x^2}+2B(x, y)\frac{\partial^2 u}{\partial x\,\partial y}+C(x, y)\frac{\partial^2 u}{\partial y^2}+$$
$$+a(x, y)\frac{\partial u}{\partial x}+b(x, y)\frac{\partial u}{\partial y}+c(x, y)u. \tag{12}$$

Отыскание решения $u=u(x, y)$ уравнения (11), удовлетворяющего начальным условиям

$$u(x, y_0)=\varphi(x), \quad u_y(x, y_0)=\varphi_1(x), \tag{13}$$

называется *задачей Коши*, а сами условия носят название *начальных данных Коши*.

Задача Коши допускает простую геометрическую интерпретацию (рис. 53): требуется найти интегральную поверхность $u=u(x, y)$ уравнения (11), проходящую через данную пространственную кривую

$$y=y_0, \quad u=\varphi(x) \tag{Г}$$

и касающуюся в точках $M(x, y_0, u)$ этой кривой заданной системы векторов $\boldsymbol{a}$, расположенных в плоскостях $x=\text{const}$ и составляющих с осью Oy угол β, определяемый равенством

$$\operatorname{tg}\beta=\varphi_1(x).$$

Если рассматривать y как время, то задача Коши имеет следующую механическую трактовку: в начальный момент времени $y=y_0$ заданы форма плоской линии $u=\varphi(x, y_0)$ и распределение скоростей ее точек $\frac{\partial u}{\partial y}=\varphi_1(x, y_0)$. Предполагая, что каждая точка $M(x, u)$ линии движется параллельно оси Ou, причем дифференциальный закон движения дается уравнением (11), требуется определить форму линии для последующих моментов времени $y>y_0$ (рис. 54).

Условия (13) задают начальные данные Коши на прямой $y = y_0$. Однако это не является обязательным: можно задавать начальные

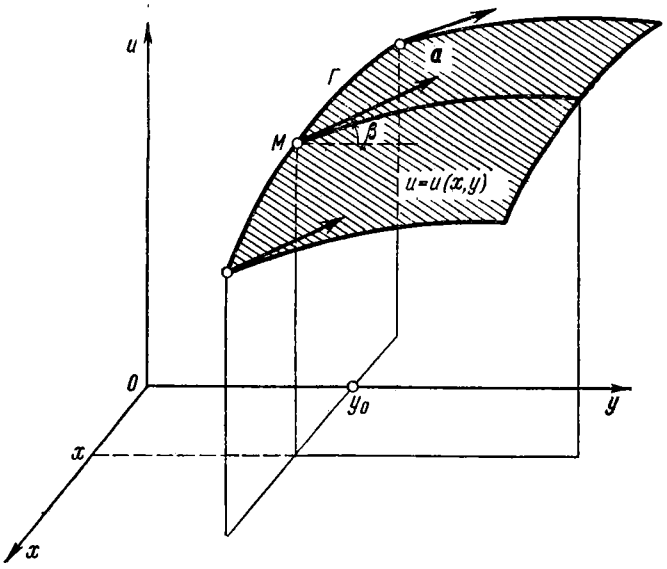

Рис. 53.

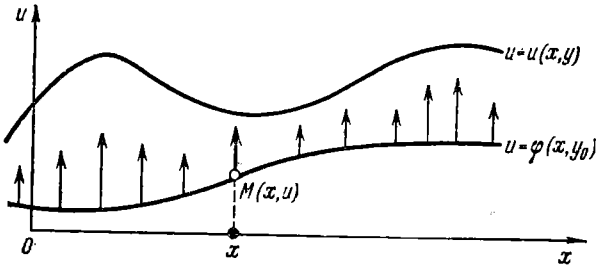

Рис. 54.

данные на любой гладкой кривой

$$\Phi(x, y) = 0. \qquad (\gamma)$$

Таким образом, приходим к *общей задаче Коши* — найти решение

$$u = u(x, y) \qquad (14)$$

дифференциального уравнения (11), удовлетворяющее начальным условиям

$$u|_\gamma = \varphi(x, y), \quad \left.\frac{\partial u}{\partial x}\right|_\gamma = \varphi_1(x, y). \qquad (15)$$

Вместо производной $\frac{\partial u}{\partial x}$ можно задавать производную $\frac{\partial u}{\partial y}$, так как на кривой γ имеем

$$\frac{\partial u}{\partial x}dx+\frac{\partial u}{\partial y}dy=d\varphi(x,\ y),\quad \frac{\partial \Phi}{\partial x}dx+\frac{\partial \Phi}{\partial y}dy=0. \tag{16}$$

Можно также задавать нормальную производную

$$\frac{\partial u}{\partial n}=\frac{\partial u}{\partial x}\cos(n,\ x)+\frac{\partial u}{\partial y}\cos(n,\ y).$$

Задача Коши обычно ставится для линейного уравнения (11) гиперболического и параболического типов.

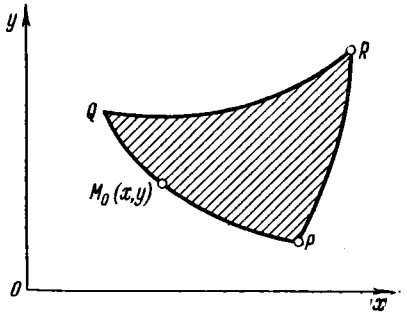

Рис. 55.

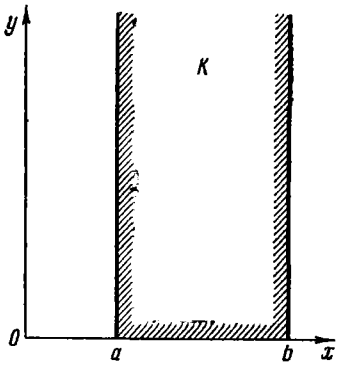

Рис. 56.

Если уравнение (11) гиперболического типа, то для единственности решения задачи Коши необходимо, чтобы начальная кривая γ не являлась характеристикой [1]. Если это последнее условие выполнено и начальные данные заданы на конечной дуге PQ кривой γ, то решение задачи Коши, вообще говоря, определено и однозначно в криволинейном треугольнике PQR (*область распространения*), образованном дугой PQ и дугами характеристик PR и QR различных семейств, проходящих через концы P и Q (рис. 55). Предполагается, конечно, что коэффициенты дифференциального уравнения определены и непрерывны в соответствующей области.

Пусть начальные данные Коши для уравнения (11) заданы на отрезке $a\leqslant x\leqslant b$, а решение $u=u(x,\ y)$ этого уравнения надо определить в полуполосе $K\{a\leqslant x\leqslant b;\ 0\leqslant y<\infty\}$ (рис. 56). Тогда для однозначности этого решения дополнительно нужно задать условия на прямых $x=a$ и $x=b$, что приводит к *смешанной задаче*. Достаточно общей задачей этого типа является нахождение

в полуполосе K решения $u=u(x, y)$ дифференциального уравнения (11), удовлетворяющего начальным и граничным условиям:

$$u(x, 0)=\varphi(x), \quad u_y(x, 0)=\varphi_1(x) \quad (a \leqslant x \leqslant b, \ y=0) \tag{17}$$

и

$$\left.\begin{aligned} &\alpha_0 u(a, y)+\alpha_1 u_x(a, y)=\psi(y), \\ &\beta_0 u(b, y)+\beta_1 u_x(b, y)=\psi_1(y), \end{aligned}\right\} \tag{18}$$

$$|\alpha_0|+|\alpha_1| \neq 0; \ |\beta_0|+|\beta_1| \neq 0, \quad 0<y<\infty.$$

Особого внимания заслуживает предельный случай, когда $a=-\infty$ или $b=\infty$. Здесь краевые условия (18) или совсем отпадают, или заменяются некоторыми условиями «на бесконечности».

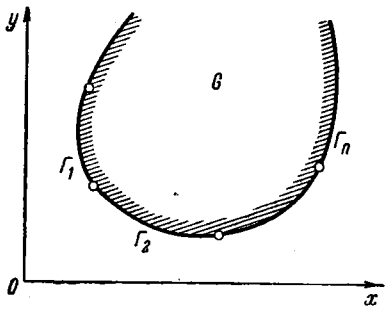

Рис. 57.

Смешанная задача для уравнения (11) в общем аспекте может быть сформулирована следующим образом: дана конечная или бесконечная область G в плоскости Oxy, имеющая кусочно-гладкую границу Γ (рис. 57). Требуется найти в области G решение дифференциального уравнения

$$L[u]=F(x, y), \tag{19}$$

если на некоторых частях границы $\Gamma_1, \Gamma_2, \ldots, \Gamma_n$ выполнены соотношения

$$\left.\begin{aligned} &L_{ij}[u]=\varphi_{ij}(x, y), \\ &(x, y) \in \Gamma_i \ (i=1, 2, \ldots, n; \ j=1, 2, \ldots, \rho_i), \end{aligned}\right\} \tag{20}$$

где L_{ij} — или *дифференциальные операторы* по переменным x и y порядка не выше первого, или *конечные соотношения*, а $\varphi_{ij}(x, y)$ — заданные функции. Задачу Коши, очевидно, можно рассматривать как частный случай этой общей смешанной задачи.

При рассмотрении физических проблем функции $\varphi_{ij}(x, y)$ обычно определяются приближенно из опыта. Поэтому решение такой смешанной задачи имеет практическую ценность лишь в том случае, если небольшие ошибки в начальных и краевых условиях не могут привести к большим отклонениям соответствующего решения. В этом случае говорят, что смешанная задача поставлена *корректно*, или, иначе говоря, непрерывно зависит от начальных и краевых условий.

Определение. Смешанная задача (19)—(20) называется *корректно поставленной* в области G, если для любого $\varepsilon>0$ можно указать число $\eta=\eta(\varepsilon)>0$ такое, что при изменении функций $\varphi_{ij}(x, y)$ на величины, модуль которых меньше, чем η, решение

$u=u(x, y)$ изменяется во всей области G меньше чем на ε. В противном случае считается, что задача *поставлена некорректно*.

Для уравнений эллиптического типа задача Коши обычно не рассматривается. Это объясняется тем, что, как правило, задача Коши для уравнений эллиптического типа поставлена некорректно, т. е. ничтожно малые изменения начальных данных могут повлечь существенное изменение решения. Покажем это на примере, идея которого принадлежит французскому математику Адамару.

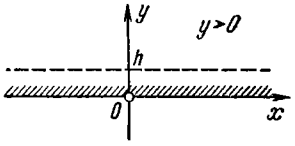

Рис. 58.

Пример 3. Пусть требуется найти в верхней полуплоскости $y>0$ (рис. 58) решение уравнения Лапласа

$$\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}=0, \tag{21}$$

удовлетворяющее при $y=0$ начальным условиям Коши:

$$u(x, 0)=0, \quad u_y(x, 0)=\frac{1}{n}\cos nx, \tag{22}$$

где n — натуральное число.

Легко проверить, что решение этой задачи есть

$$u(x, y)=\frac{1}{n^2}\cos nx \operatorname{sh} ny. \tag{23}$$

Хотя при достаточно большом n начальные данные (22) могут быть сделаны сколь угодно малыми:

$$|u(x, 0)|<\eta, \quad |u_y(x, 0)|<\eta,$$

тем не менее решение (23) для любого $\eta>0$ не ограничено при $n\to\infty$ в каждой полосе $0<y<h$. Действительно, если $y>0$, то, например, имеем

$$u(0, y)=\frac{1}{n^2}\operatorname{sh} ny \to \infty \quad \text{при } n\to\infty.$$

Если же взять чисто нулевые начальные условия $u(x, 0)=0$, $u_y(x, 0)=0$, то решение $u_0(x, y)$ этой последней задачи Коши, очевидно, будет

$$u_0(x, y)\equiv 0.$$

Таким образом, здесь сколь угодно малое изменение начальных условий приводит к неограниченно большому расхождению решений.

§ 3. Краевые задачи для уравнений эллиптического типа

Исследования стационарных процессов различной физической природы (колебания, теплопроводность и др.) часто приводят к уравнениям эллиптического типа

$$L[u]\equiv\Delta u+au_x+bu_y+cu=F(x, y), \tag{1}$$

где $a=a(x, y)$, $b=b(x, y)$, $c=c(x, y)$ и $F(x, y)$ — непрерывные функции. Для этих уравнений обычно ставятся лишь краевые задачи, так как задача Коши, как было показано в предыдущем параграфе, для уравнений эллиптического типа может быть некорректной.

Наиболее часто встречаются следующие *краевые задачи*.

Первая краевая задача. На контуре Г, ограничивающем область G (рис. 59), задана непрерывная функция $\varphi(P)=\varphi(x, y)$. Требуется найти функцию $u(P)=u(x, y)$, удовлетворяющую внутри G уравнению (1) и принимающую на границе заданные значения $\varphi(P)$, т. е. должны быть выполнены условия:

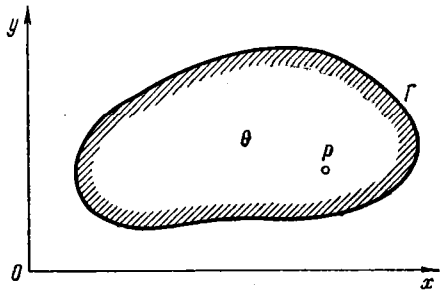

Рис. 59.

$$L[u(P)]=F(P) \text{ при } P\in G;$$
$$u(P)=\varphi(P) \text{ при } P\in\Gamma.$$

Вторая краевая задача. На контуре Г, ограничивающем область G, задана непрерывная функция $\varphi_1(P)$. Требуется найти функцию $u(P)=u(x, y)$, удовлетворяющую внутри G уравнению (1), нормальная производная которой на Г принимает заданные значения $\varphi_1(P)$, т. е. требуется, чтобы

$$L[u(P)]=F(P) \text{ при } P\in G, \quad \frac{\partial u(P)}{\partial n}=\varphi_1(P) \text{ при } P\in\Gamma.$$

Третья краевая задача. На контуре Г, ограничивающем область G (рис. 59), задана непрерывная функция $\psi(P)=\psi(x, y)$. Требуется найти функцию $u(P)=u(x, y)$ такую, чтобы

$$L[u(P)]=F(P) \text{ при } P\in G,$$
$$\alpha_0 u(P)+\alpha_1\frac{\partial u(P)}{\partial n}=\psi(P) \text{ при } P\in\Gamma,$$

где $|\alpha_0|+|\alpha_1|\neq 0$.

Третью краевую задачу можно рассматривать как общую. Действительно, при $\alpha_0=1$ и $\alpha_1=0$ получаем первую краевую задачу, а при $\alpha_0=0$ и $\alpha_1=1$ — вторую краевую задачу. Заметим, что если область G ограниченная, то соответствующая краевая задача называется *внутренней*, в противном случае — *внешней*.

Для уравнения Лапласа $\Delta u=0$ первая краевая задача называется *задачей Дирихле*, вторая — *задачей Неймана* и третья — *смешанной краевой задачей*.

§ 4. Некоторые сведения о гармонических функциях. Единственность решения задачи Дирихле

Определение. Функция $u(x, y)$, имеющая непрерывные частные производные второго порядка в области G и удовлетворяющая внутри G уравнению Лапласа, называется *гармонической функцией*.

Простейшими гармоническими функциями двух переменных x и y являются: линейная функция $u = ax + by + c$; функция вида $u = \ln r$, где $r = \sqrt{(x-x_0)^2 + (y-y_0)^2}$ (основное решение уравнения Лапласа).

Задача Дирихле в иных терминах может быть сформулирована следующим образом: найти функцию, непрерывную в данной замкнутой области $\overline{G} = G + \Gamma$, гармоническую в области G и принимающую на ее границе Γ непрерывные заданные значения.

Единственность решения задачи Дирихле и непрерывная зависимость ее от краевых условий (корректность краевой задачи) вытекают из следующих свойств гармонических функций.

Свойство I *(принцип максимума)*. Гармоническая в ограниченной области функция, непрерывная в замкнутой области $\overline{G} = G + \Gamma$, не может принимать внутри этой области значений больших, чем максимум ее значений на границе Γ, и меньших, чем минимум ее значений на Γ.

Доказательство. Пусть M—максимум значений $u(x, y)$ на границе Γ. Допустим, что функция $u(x, y)$ в некоторой точке $P_0(x_0, y_0)$ внутри G принимает значение $\mu = u(x_0, y_0)$, причем $\mu > M$.

Составим вспомогательную функцию

$$v(x, y) = u(x, y) + \frac{\mu - M}{2d^2}\left[(x-x_0)^2 + (y-y_0)^2\right],$$

где d—диаметр области G. Очевидно, имеем

$$v(x_0, y_0) = u(x_0, y_0) = \mu,$$

причем при $(x, y) \in \Gamma$ выполнено неравенство

$$v(x, y) \leqslant M + \frac{\mu - M}{2d^2} d^2 = \frac{M + \mu}{2} < \mu.$$

Следовательно, функция $v(x, y)$ достигает своего наибольшего значения внутри области G в некоторой точке $(\overline{x}, \overline{y})$, причем в этой точке будут выполнены необходимые условия для максимума функции:

$$\frac{\partial v}{\partial x} = \frac{\partial v}{\partial y} = 0, \quad \frac{\partial^2 v}{\partial x^2} \leqslant 0, \quad \frac{\partial^2 v}{\partial y^2} \leqslant 0.$$

Из соотношения

$$\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{2(\mu - M)}{d^2} = 0 + \frac{2(\mu - M)}{d^2} > 0$$

вытекает, что по меньшей мере одна из производных $\frac{\partial^2 v}{\partial x^2}$ или $\frac{\partial^2 v}{\partial y^2}$ положительна внутри G. Поэтому функция $v(x, y)$ ни в какой внутренней точке области G не может иметь максимума, и, следовательно, приходим к противоречию. Таким образом, $u(x, y) \leqslant M$.

Аналогично доказывается, что $u(x, y) \geqslant m$, где m — наименьшее значение функции $u(x, y)$ на границе Γ.

Следствие. Пусть функция $u = u(x, y)$ — гармоническая в ограниченной области G и непрерывная в замкнутой области $\overline{G} = G + \Gamma$. В таком случае справедливо неравенство

$$\underline{u} \leqslant u(x, y) \leqslant \overline{u},$$

где $\underline{u} = \min u(x, y)$ на Γ, $\overline{u} = \max u(x, y)$ на Γ.

Замечание. Можно доказать более сильное утверждение, что гармоническая в ограниченной и замкнутой области G функция, отличная от константы, не принимает внутри G наибольшего и наименьшего значений.

Свойство II *(единственность задачи Дирихле)*. Задача Дирихле для замкнутой и ограниченной области может иметь лишь единственное решение, т. е. не существует двух непрерывных гармонических функций в замкнутой ограниченной области G, принимающих на границе одни и те же значения.

Доказательство. Допустим, что две функции $u_1(x, y)$ и $u_2(x, y)$, гармонические в области G, совпадают всюду на ее границе. Рассмотрим функцию

$$u(x, y) = u_1(x, y) - u_2(x, y).$$

Очевидно, что $u(x, y)$ — гармоническая функция, обращающаяся в нуль на границе. По свойству I эта функция не может принимать внутри G значений больше или меньше нуля, следовательно, $u(x) \equiv 0$ внутри G и $u_1(x) \equiv u_2(x)$.

Замечание. Из свойства II не следует, что задача Дирихле для ограниченной замкнутой области G имеет решение; это свойство лишь утверждает, что если существует решение задачи Дирихле для области G, то оно единственно.

Можно доказать [20], что если область G выпуклая, т. е. вместе с двумя своими точками содержит соединяющий их отрезок, и граница ее Γ кусочно-гладкая, то задача Дирихле для такой области с непрерывными данными на ее границе Γ действительно имеет решение (теорема Неймана).

Свойство III *(корректность задачи Дирихле)*. Решение задачи Дирихле для замкнутой и ограниченной области непрерывно зависит от граничных данных.

Доказательство. Допустим, что $u_1(x, y)$ и $u_2(x, y)$ — решения задачи Дирихле, соответственно принимающие на границе значения $\varphi_1(x, y)$ и $\varphi_2(x, y)$.

Пусть всюду на границе Γ выполнено неравенство

$$|\varphi_1(x, y) - \varphi_2(x, y)| < \varepsilon,$$

где ε — произвольно малое положительное число.

Рассмотрим гармоническую функцию

$$u(x, y) = u_1(x, y) - u_2(x, y).$$

На границе Γ эта функция принимает значение

$$\varphi(x, y) = \varphi_1(x, y) - \varphi_2(x, y).$$

Так как $-\varepsilon < \varphi(x, y) < \varepsilon$ на Γ, то по свойству I имеем

$$-\varepsilon < u(x, y) < \varepsilon \quad \text{при } (x, y) \in G,$$

т. е. $-\varepsilon < u_1(x, y) - u_2(x, y) < \varepsilon$ или $|u_1(x, y) - u_2(x, y)| < \varepsilon$.

Таким образом, для задачи Дирихле требование корректности выполнено при $\eta = \varepsilon$ (см. § 3).

§ 5. Уравнение Лапласа в конечных разностях

Для получения конечно-разностного уравнения, соответствующего уравнению Лапласа

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0, \tag{1}$$

достаточно, выбрав шаг $h > 0$, заменить производные $\frac{\partial^2 u}{\partial x^2}$ и $\frac{\partial^2 u}{\partial y^2}$ отношениями конечных разностей по формулам

$$\frac{\partial^2 u}{\partial x^2} \approx \frac{u(x+h, y) - 2u(x, y) + u(x-h, y)}{h^2},$$

$$\frac{\partial^2 u}{\partial y^2} \approx \frac{u(x, y+h) - 2u(x, y) + u(x, y-h)}{h^2}.$$

Тогда будем иметь

$$\frac{u(x+h, y) - 2u(x, y) + u(x-h, y)}{h^2} + \frac{u(x, y+h) - 2u(x, y) + u(x, y-h)}{h^2} = 0,$$

и, следовательно,

$$u(x, y) = \frac{1}{4}[u(x+h, y) + u(x-h, y) + u(x, y+h) + u(x, y-h)]. \tag{2}$$

Однако, чтобы иметь возможность оценить точность такой замены, следует идти по несколько иному пути, используя для получения конечно-разностного уравнения формулу Тейлора

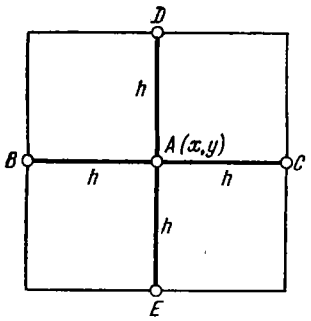

Рис. 60.

$$f(x+h,\ y+k)=$$
$$=f(x,\ y)+\left(\frac{\partial}{\partial x}h+\frac{\partial}{\partial y}k\right)f(x,\ y)+$$
$$+\frac{1}{2!}\left(\frac{\partial}{\partial x}h+\frac{\partial}{\partial y}k\right)^2 f(x,\ y)+\ldots$$
$$\ldots+\frac{1}{n!}\left(\frac{\partial}{\partial x}h+\frac{\partial}{\partial y}k\right)^n f(x+\Theta h, y+\Theta k), \tag{3}$$

где $0<\Theta<1$.

При этом пользуются различными схемами. Рассмотрим две основные схемы.

Первая основная схема. Рассмотрим точки $A(x,\ y)$, $B(x-h,\ y)$, $C(x+h,\ y)$, $D(x,\ y+h)$, $E(x,\ y-h)$, лежащие в центре квадрата и на серединах его сторон (рис. 60), и выразим значения функции u в точках B, C, D, E через значения этой функции и ее производных в центральной точке квадрата A $(x,\ y)$.

Согласно формуле (3), полагая в ней $n=4$, имеем:

$$\left.\begin{aligned}
&u(x-h,\ y)=\\
&\qquad=u(x,\ y)-hu_x+\frac{1}{2!}h^2u_{xx}-\frac{1}{3!}h^3u_{xxx}+\frac{1}{4!}h^4\bar{u}_{xxxx},\\
&u(x+h,\ y)=\\
&\qquad=u(x,\ y)+hu_x+\frac{1}{2!}h^2u_{xx}+\frac{1}{3!}h^3u_{xxx}+\frac{1}{4!}h^4\bar{\bar{u}}_{xxxx},\\
&u(x,\ y-h)=\\
&\qquad=u(x,\ y)-hu_y+\frac{1}{2!}h^2u_{yy}-\frac{1}{3!}h^3u_{yyy}+\frac{1}{4!}h^4\tilde{u}_{yyyy},\\
&u(x,\ y+h)=\\
&\qquad=u(x,\ y)+hu_y+\frac{1}{2!}h^2u_{yy}+\frac{1}{3!}h^3u_{yyy}+\frac{1}{4!}h^4\tilde{\tilde{u}}_{yyyy},
\end{aligned}\right\} \tag{4}$$

где u_x, u_y, u_{xx}, u_{yy}, u_{xxx}, u_{yyy} — значения производных в точке $A(x,\ y)$, а $\bar{u}_{xxxx}$, $\bar{\bar{u}}_{xxxx}$, $\tilde{u}_{yyyy}$, $\tilde{\tilde{u}}_{yyyy}$ — производные в некоторых промежуточных точках.

Складывая равенства (4), получаем

$$u(x-h,\ y)+u(x+h,\ y)+u(x,\ y-h)+u(x,\ y+h)=$$
$$=4u(x,\ y)+h^2(u_{xx}+u_{yy})+R_h(x,\ y), \tag{5}$$

где *остаточный член*

$$R_h(x, y) = \frac{h^4}{4!}\left[\bar{u}_{xxxx} + \bar{\bar{u}}_{xxxx} + \tilde{u}_{yyyy} + \tilde{\tilde{u}}_{yyyy}\right]$$

при $u \in C^{(4)}$ имеет порядок $O(h^4)$. Отсюда будем иметь

$$u(x-h, y) + u(x+h, y) + u(x, y-h) + u(x, y+h) = \\ = 4u(x, y) + h^2\Delta u + O(h^4),$$

и, следовательно,

$$\Delta u = \frac{1}{h^2}\left[u(x-h, y) + u(x+h, y) + u(x, y-h) + \right. \\ \left. + u(x, y+h) - 4u(x, y)\right] + O(h^2). \qquad (6)$$

Формула (6) выражает оператор Лапласа Δu через конечные разности и называется первой *основной конечно-разностной формой*

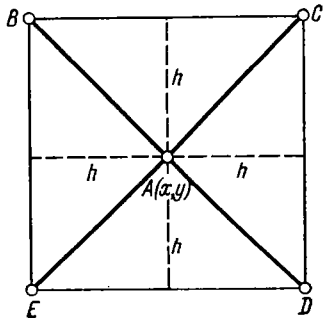

Рис. 61.

оператора Лапласа. Откидывая в уравнении (6) член $O(h^2)$, получим, что уравнению Лапласа $\Delta u = 0$ приближенно соответствует следующее уравнение в конечных разностях:

$$u(x, y) = \frac{1}{4}\left[u(x+h, y) + u(x-h, y) + \right. \\ \left. + u(x, y+h) + u(x, y-h)\right], \qquad (7)$$

что совпадает с уравнением (2).

Вторая основная схема. Рассмотрим точки $A(x, y)$, $B(x-h, y+h)$, $C(x+h, y+h)$, $D(x+h, y-h)$, $E(x-h, y-h)$, лежащие в центре и вершинах квадрата (рис. 61).

Как и в первой схеме, выразим значения функции u в точках B, C, D, E через значения этой функции и ее производных в точке A.

Полагая $n=4$ в формуле (3), получим:

$$\left.\begin{aligned}
u(x+h,\ y-h) &= u(x,\ y)+h(u_x-u_y)+\\
&\quad+\frac{h^2}{2!}(u_{xx}-2u_{xy}+u_{yy})+\\
&\quad+\frac{h^3}{3!}(u_{xxx}-3u_{xxy}+3u_{xyy}-u_{yyy})+\\
&\quad+\frac{h^4}{4!}\left(\frac{\partial}{\partial x}-\frac{\partial}{\partial y}\right)^4 u(\xi_1,\ \eta_1),\\
u(x-h,\ y-h) &= u(x,\ y)+h(-u_x-u_y)+\\
&\quad+\frac{h^2}{2!}(u_{xx}+2u_{xy}+u_{yy})+\\
&\quad+\frac{h^3}{3!}(-u_{xxx}-3u_{xxy}-3u_{xyy}-u_{yyy})+\\
&\quad+\frac{h^4}{4!}\left(-\frac{\partial}{\partial x}-\frac{\partial}{\partial y}\right)^4 u(\xi_2,\ \eta_2),\\
u(x-h,\ y+h) &= u(x,\ y)+h(-u_x+u_y)+\\
&\quad+\frac{h^2}{2!}(u_{xx}-2u_{xy}+u_{yy})+\\
&\quad+\frac{h^3}{3!}(-u_{xxx}+3u_{xxy}-3u_{xyy}+u_{yyy})+\\
&\quad+\frac{h^4}{4!}\left(-\frac{\partial}{\partial x}+\frac{\partial}{\partial y}\right)^4 u(\xi_3,\ \eta_3),\\
u(x+h,\ y+h) &= u(x,\ y)+h(u_x+u_y)+\\
&\quad+\frac{h^2}{2!}(u_{xx}+2u_{xy}+u_{yy})+\\
&\quad+\frac{h^3}{3!}(u_{xxx}+3u_{xxy}+3u_{xyy}+u_{yyy})+\\
&\quad+\frac{h^4}{4!}\left(\frac{\partial}{\partial x}+\frac{\partial}{\partial y}\right)^4 u(\xi_4,\ \eta_4).
\end{aligned}\right\}\qquad(8)$$

Складывая равенства (8), будем иметь

$$u(x+h,\ y-h)+u(x-h,\ y-h)+u(x-h,\ y+h)+$$
$$+u(x+h,\ y+h)=4u(x,\ y)+2h^2\Delta u+O(h^4),$$

откуда

$$\Delta u=\frac{1}{2h^2}[u(x+h,\ y-h)+u(x-h,\ y-h)+$$
$$+u(x+h,\ y+h)+u(x-h,\ y+h)-4u(x,\ y)]+O(h^2).$$

Откидывая остаточный член $O(h^2)$, получаем, что уравнение Лапласа $\Delta u=0$ приближенно можно заменить конечно-разностным уравнением

$$u(x,\ y)=\frac{1}{4}[u(x+h,\ y-h)+u(x-h,\ y-h)+$$
$$+u(x-h,\ y+h)+u(x+h,\ y+h)].\qquad(9)$$

§ 6. Решение задачи Дирихле методом сеток

Идея *метода сеток* (или, иначе, метода *конечных разностей*) для приближенного решения краевых задач для двумерных дифференциальных уравнений заключается в следующем:

1) в плоской области G, в которой разыскивается решение, строится *сеточная область* G_h, состоящая из одинаковых ячеек (рис. 62) и приближающая данную область G;

2) заданное дифференциальное уравнение заменяется в узлах построенной сетки соответствующим конечно-разностным уравнением;

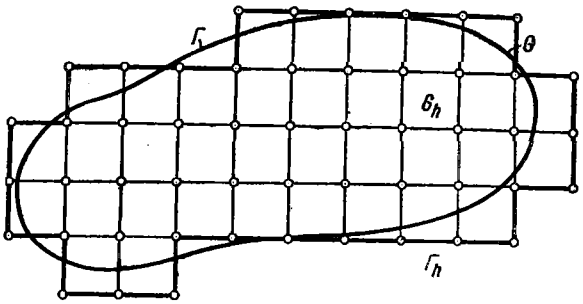

Рис. 62.

3) на основании граничных условий устанавливаются значения искомого решения в граничных узлах области G_h.

Решив полученную систему конечно-разностных уравнений, для чего, вообще говоря, нужно решить алгебраическую систему с большим числом неизвестных, мы найдем значения искомой функции в узлах сетки, т. е. будем иметь численное решение нашей задачи.

Выбор сеточной области производится в зависимости от конкретной задачи, но во всех случаях контур Γ_h сеточной области G_h следует выбирать так, чтобы он возможно лучше аппроксимировал контур Γ заданной области G.

Сеточная область может состоять из квадратных, прямоугольных, треугольных и других клеток. От выбора основного размера клетки h зависит величина остаточного члена R_h при замене дифференциального уравнения конечно-разностным. Следовательно, размер h теоретически должен определяться требованием, чтобы этот остаточный член был меньше погрешности, допустимой при решении. Однако такой путь не всегда целесообразен, так как получаемый при этом размер h настолько мал и, следовательно, число клеток настолько велико, что решение оказывается практически невыполнимым.

Обычно задача решается сначала при большом значении h, т. е. при малом числе клеток, и лишь после того, как задача грубо приближенно решена для этой крупной сетки, переходят к более мелкой сетке или во всей рассматриваемой области, или в какой-нибудь ее части.

Идея метода сеток известна давно и восходит еще к Эйлеру. Однако практическое использование этого метода наталкивалось на серьезные трудности, так как получение с его помощью достаточно точного решения краевой задачи обычно приводило к колоссальным системам алгебраических уравнений, на решение которых при ручном счете требовались годы вычислительного труда. Положение резко изменилось с появлением быстродействующих электронных вычислительных машин. Метод сеток допускает удобную реализацию на электронных счетных машинах, так как применение его обычно сводится к массовой повторяемости однородных циклов. В настоящее время метод сеток является одним из наиболее эффективных методов решения линейных, а также отчасти нелинейных задач математической физики [15].

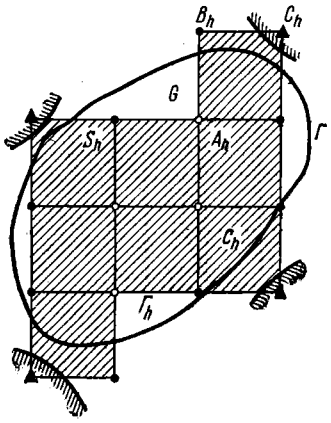

Рис. 63.

Покажем применение метода сеток для построения решения задачи Дирихле

$$\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}=0 \text{ при } (x, y)\in G \text{ и } u(P)=\varphi(P) \text{ при } P\in\Gamma, \qquad (1)$$

где $\varphi(P)=\varphi(x, y)$ — заданная непрерывная функция, причем для простоты рассмотрим лишь случай квадратной сетки. Будем предполагать, что область G ограничена простым замкнутым кусочно-гладким контуром Γ.

Выбрав шаг h, построим квадратную сетку

$$x_i=x_0+ih,\ y_j=y_0+jh\ (S_h)$$
$$(i,\ j=0,\ \pm1,\ \pm2,\ \ldots)$$

с таким расчетом, чтобы узлы $(x_i,\ y_j)$ сетки S_h или принадлежали области G, или отстояли от ее границы Γ на расстоянии меньшем, чем h.

Точки (узлы) сетки S_h называются *соседними*, если они удалены друг от друга в направлении оси Ox или оси Oy на расстояние, равное шагу сетки h. Узел A_h сетки S_h называется *внутренним*,

если он принадлежит области G, а все четыре соседних с ним узла — множеству S_h; в противном случае он называется *граничным* (например, узлы B_h и C_h сетки S_h) (на рис. 63 внутренние узлы обозначены светлыми кружками, а граничные — темными кружками и темными треугольниками).

Граничный узел сетки S_h называется узлом *первого рода*, если он имеет соседний внутренний узел этой сетки (например, узел B_h на рис. 63); в противном случае граничный узел называется узлом *второго рода* (узел C_h на рис. 63). Внутренние узлы и граничные узлы первого рода сетки S_h называются *расчетными точками*. Граничные узлы второго рода не входят в вычисление и могут быть изъяты из сетки S_h (на рис. 63 граничные узлы второго рода обозначены темными треугольниками).

Относительно сетки S_h предположим, что множество ее расчетных точек «связное», т. е. любые две расчетные точки можно соединить цепочкой узлов, каждые два смежных элемента которой являются соседними узлами. Кроме того, будем считать многоугольную сеточную область G_h выбранной так, чтобы ее геометрическая граница Γ_h возможно ближе примыкала к границе Γ области G. Заметим, что узловые точки контура Γ_h могут лежать как внутри, так и вне области G.

Значение искомой функции $u=u(x, y)$ в точках (x_i, y_j) обозначим через $u_{ij}=u(x_i, y_j)$. Следуя общей схеме (см. § 5 (2)), для каждой внутренней точки (x_i, y_j) сетки S_h заменяем дифференциальное уравнение (1) конечно-разностным уравнением

$$u_{ij}=\frac{1}{4}(u_{i-1,j}+u_{i+1,j}+u_{i,j-1}+u_{i,j+1}), \tag{2}$$

где $(x_{i\pm1}, y_{j\pm1})$ — расчетные точки.

В граничных узлах первого рода B_h сетки S_h полагаем

$$u(B_h)=u(B)=\varphi(B), \tag{3}$$

где B — ближайшая к B_h точка границы Γ.

Система (2) является неоднородной линейной системой, причем число неизвестных (т. е. число внутренних узлов сетки) равно числу уравнений. Система (2) всегда совместна и имеет единственное решение. Чтобы доказать это, достаточно убедиться в том, что соответствующая однородная система имеет лишь нулевое решение. Однородная система, очевидно, формально может быть записана в виде системы (2), с той лишь разницей, что значение функции $\varphi(P)$ на границе Γ следует положить тождественно равным нулю: $\varphi(P)\equiv 0$.

Однородная система (2) всегда совместна, так как эта система имеет тривиальное решение $u_{ij}\equiv 0$. Покажем, что однородная система (2) не может иметь решений $u_{ij}\not\equiv 0$. Пусть, например, для некоторого решения одно из ее неизвестных $u_{pq}\neq 0$. Для

определенности будем считать $u_{pq} > 0$. Обозначим через u_{PQ} наибольшую компоненту рассматриваемого решения, т. е. положим

$$u_{PQ} \geqslant u_{ij} \tag{4}$$

для всех узлов сетки S_h. В силу неравенства (4) будем иметь

$$u_{PQ} \geqslant u_{pq} > 0. \tag{5}$$

На основании системы (2) получаем

$$u_{PQ} = \frac{1}{4}(u_{P-1,\,Q} + u_{P+1,\,Q} + u_{P,\,Q-1} + u_{P,\,Q+1}). \tag{6}$$

Учитывая неравенство (4), заключаем, что

$$u_{P-1,\,Q} \leqslant u_{PQ},\ u_{P+1,\,Q} \leqslant u_{PQ},\ u_{P,\,Q-1} \leqslant u_{PQ},\ u_{P,\,Q+1} \leqslant u_{PQ}.$$

Ни одно из последних четырех неравенств не является строгим, так как если бы это имело место, то, складывая все четыре неравенства и учитывая формулу (6), мы получили бы $u_{PQ} < u_{PQ}$ (?!). Поэтому

$$u_{P-1,\,Q} = u_{P+1,\,Q} = u_{P,\,Q-1} = u_{P,\,Q+1} = u_{PQ}. \tag{7}$$

Проводя аналогичные рассуждения для точек $u_{P+1,\,Q}$, $u_{P+2,\,Q}, \ldots$, на конечном шаге достигнем некоторого граничного узла $u_{P+m,\,Q}$ сетки S_h с ближайшей точкой $M \in \Gamma$, где положено

$$u_{P+m,\,Q} = \varphi(M) = 0.$$

Таким образом, из цепи равенств (7) имеем $u_{PQ} = 0$, что противоречит неравенству (5).

Итак, однородная система (2) не может иметь положительных решений. Аналогично доказывается, что эта система не может иметь отрицательных решений. Следовательно, $u_{ij} \equiv 0$ для каждого решения, и, значит, неоднородная система (2) совместна и имеет единственное решение.

Решив систему (2), получим приближенные значения искомой функции $u = u(x, y)$ в узлах сеточной области G_h. Тем самым будет найдено приближенное численное решение задачи Дирихле для области G_h. Можно показать, что в общем случае погрешность приближенного решения имеет порядок $O(h)$.

§ 7. Процесс Либмана

Если число узлов сетки S_h велико, то непосредственное решение системы (2) из § 6 становится затруднительным. Кроме того, для криволинейной области G значения функции u в граничных узлах сетки S_h выбраны слишком грубо. Эти обстоятельства за-

ставляют для решения указанной системы прибегать к итерационным методам с одновременным исправлением граничных значений.

Согласно *процессу усреднения Либмана*, выбрав начальные приближения $u_{ij}^{(0)}$, последовательные приближения $u_{ij}^{(k)}$ для внутренних узлов (x_i, y_j) сетки S_h определяем по формуле

$$u_{ij}^{(k)} = \frac{1}{4}\left[u_{i-1,j}^{(k-1)} + u_{i+1,j}^{(k-1)} + u_{i,j-1}^{(k-1)} + u_{i,j+1}^{(k-1)}\right] \quad (k = 1, 2, \ldots). \qquad (1)$$

Что касается граничных узлов A_h сетки S_h, то значения функции $u(A_h)$ в этих узлах последовательно исправляем по формулам *линейной интерполяции*:

$$u^{(0)}(A_h) = u(A) = \varphi(A),$$

$$u^{(k)}(A_h) = u(A) + \frac{u^{(k-1)}(B) - u(A)}{h+\delta}\delta$$

$$(k = 1, 2, \ldots), \qquad (2)$$

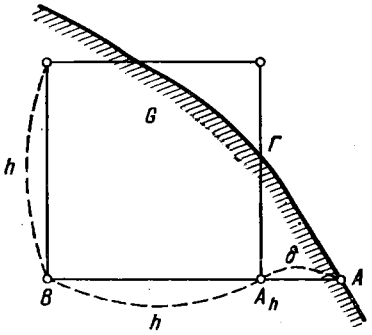

Рис. 64.

где A — ближайшая к A_h точка границы Γ $(u(A) = \varphi(A))$, B — ближайший к A_h внутренний узел сетки S_h (рис. 64) и δ — удаление узла A_h от точки A, причем $\delta > 0$, если A_h — внутренняя точка области G, и $\delta < 0$, если A_h — внешняя точка области G. В частном случае, если узел A_h лежит на границе Γ $(A_h \equiv A,\ \delta = 0)$, то имеем точно

$$u^{(k)}(A_h) = u(A) = \varphi(A).$$

На практике после некоторого шага k можно считать $u^{(k)}(A_h)$ неизменными (например, если эти значения установятся с заданной степенью точности).

За начальные значения $u_{ij}^{(0)}$ теоретически можно взять любую систему чисел. Однако следует иметь в виду, что в силу принципа максимума (см. § 4, свойство I) для значений искомой функции $u(x, y)$ должны быть выполнены неравенства

$$m \leqslant u_{ij} \leqslant M,$$

где $m = \min\limits_{\Gamma} \varphi(P)$ и $M = \max\limits_{\Gamma} \varphi(P)$. Поэтому разумно полагать

$$m \leqslant u_{ij}^{(0)} \leqslant M.$$

Практически для выбора $u_{ij}^{(0)}$ грубо решают задачу Дирихле в области G с помощью крупной сетки, а затем найденные значения используют для решения задачи Дирихле на данной мелкой сетке. Для начала процесса обычно применяют линейную интерполяцию.

Доказывается [1], [3], [6], что для любого шага сетки h процесс Либмана независимо от выбора начальных значений сходится, т. е. существует

$$\lim_{k\to\infty} u_{ij}^{(k)} = u_{ij},$$

причем погрешность приближенного решения имеет порядок $O(h^2)$.

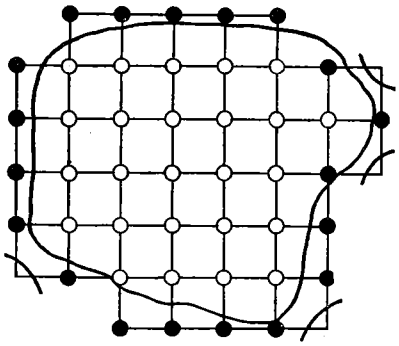

Рис. 65.

Для практического проведения вычислений по методу итерации полезно приготовить достаточное число специальных *вычислительных шаблонов* [9]. Способ заготовления этих шаблонов для случая неизменных граничных значений следующий. Пусть область G, в которой должно быть найдено решение задачи Дирихле, покрыта сеткой (рис. 65). Внутренние узлы этой сетки отмечены белыми кружками; граничные узлы, в которых известны значения искомой функции, отмечены черными кружками.

Для построения вычислительного шаблона строим вторую сетку, линии которой проходят посредине между линиями первой, причем так, что узлы первой сетки (внутренние и граничные) попадают в центры клеток второй сетки (рис. 66).

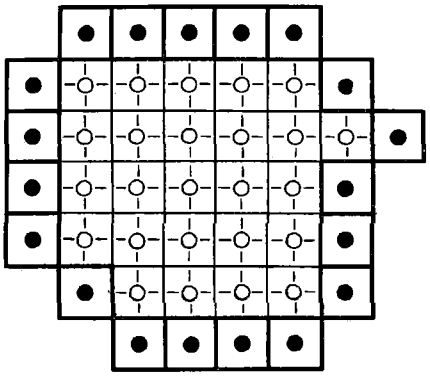

Рис. 66.

Клетки второй сетки, в центрах которых лежат граничные узлы первой сетки, обведем жирной чертой. Готовый вычислительный шаблон № 1 приведен на рис. 67.

В обведенные жирной чертой клетки шаблона № 1 вписываются неизменные граничные значения, определенные на контуре Γ_h. Внутренние клетки будут заполняться последовательно итерационным процессом. Поэтому нужно заготовить достаточное число одинаковых

шаблонов № 2, 3, ..., состоящих из одних внутренних клеток*) такого же размера, как клетки шаблона № 1 (рис. 68).

Внутренние клетки шаблона № 1 заполняем начальными значениями процесса итерации (произвольными числами или решениями задачи Дирихле, полученными для более крупной сетки). Когда шаблон № 1 заполнен, начинается заполнение шаблона № 2 таким образом, чтобы в каждой его клетке

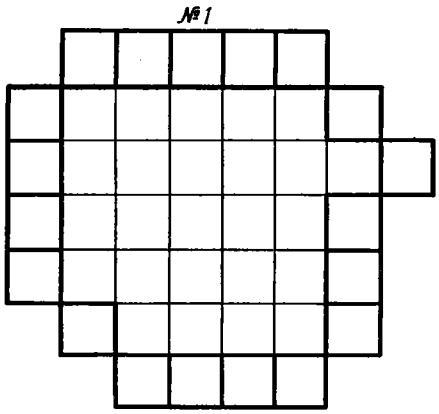

Рис. 67.

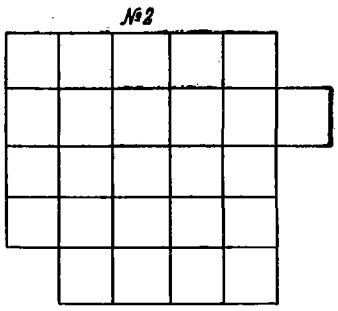

Рис. 68.

было записано среднее арифметическое четырех чисел, стоящих в соответствующих клетках шаблона № 1.

Очевидно, что при заполнении шаблона № 2 также участвуют и значения, стоящие в граничных клетках шаблона № 1.

После заполнения шаблона № 2 его накладывают на шаблон № 1, оставляя при этом открытыми граничные точки последнего, и аналогичным способом заполняют шаблон № 3, используя шаблон № 2.

Процесс продолжается до тех пор, пока в пределах заданной точности не совпадут два последних шаблона.

Пример 1. Найти приближенное решение уравнения

$$\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}=0,$$

удовлетворяющее на окружности $x^2+y^2=16$ (Г) условию

$$u(x, y)\,|_{\Gamma}=x^2y^2.$$

Решение. В силу симметричности решения рассмотрим четверть круга.

---

*) Понятно, что если граничные значения подлежат исправлению, то в шаблонах следует сохранить также и граничные клетки.

1-й этап. Берем крупную сетку с шагом $h=2$ (рис. 69). Ближайшая к узлу $A(4, 2)$ сетки точка границы Г есть $M(\sqrt{12}, 2)$, поэтому полагаем $u(A)\approx u(M)=12\cdot 2^2=48$. Аналогично для узла сетки $A'(2, 4)$ ближайшая

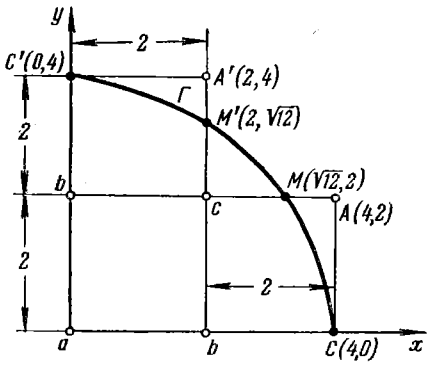

Рис. 69.

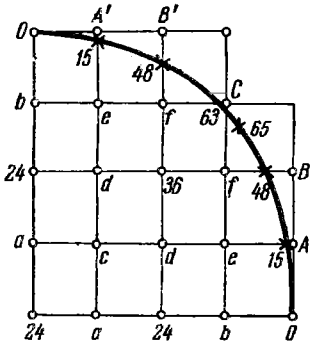

Рис. 70.

точка границы Г есть $M'(2, \sqrt{12})$, поэтому $u(A')\approx u(M')=48$. В узлах $C(4, 0)$ и $C'(0, 4)$ сетки, очевидно, имеем

$$u(C)=u(C')=0.$$

Обозначая через a, b и c значения функции u во внутренних узлах сетки (рис. 69) и учитывая симметрию задачи, составляем систему конечно-разностных уравнений

$$a=\frac{1}{4}\cdot 4b,\quad b=\frac{1}{4}(2c+a+0),\qquad c=\frac{1}{4}(48+48+2b).$$

Из этой системы находим:

$$a=24,\ b=24,\ c=36.$$

2-й этап. Берем более мелкую сетку (рис. 70) с шагом $h=1$ при неуточненных граничных значениях. Полагаем

$$u(A)=u(A')=15,\ u(B)=u(B')=48,\ u(C)=63.$$

Используя значения функции $u(x, y)$ в узлах крупной сетки с шагом $h=2$ и в граничных узлах и учитывая симметрию задачи, составляем конечно-разностные уравнения по первой и второй схеме для значений a, b, c, d, e, f искомой функции u в узлах сетки с шагом $h=1$ (рис. 70). Имеем:

$$\left.\begin{aligned} e&=\frac{1}{4}(0+36+48+24), & f&=\frac{1}{4}(48+e+63+36),\\ b&=\frac{1}{4}(e+e+0+24), & c&=\frac{1}{4}(24+24+24+36),\\ d&=\frac{1}{4}(e+c+24+36), & a&=\frac{1}{4}(24+24+c+c).\end{aligned}\right\}$$

Отсюда приближенно находим:

$$a=26,\ b=20,\ c=27,\ d=28,\ e=27,\ f=44.$$

3-й этап. Уточняем значения $u(x, y)$ в граничных узлах. Используя формулы (2) и полученные значения во внутренних узлах сетки, находим:

$$u(A)=u(A')=13,\ u(B)=u(B')=49,\ u(C)=73.$$

4-й этап. На основе полученных данных строим систему шаблонов (№ 1—7) и последовательно уточняем (с точностью до единицы) значения искомой функции $u(x, y)$ во внутренних узлах.

Шаблоны № 6 и 7 совпадают с точностью до единицы.

Отметим, что точным решением этой задачи является функция

$$u=x^2y^2+\frac{1}{8}[256-(x^2+y^2)^2].$$

Для сравнения приводим значения точного решения в узлах сетки (шаблон № 7а).

Для оценки точности решения, полученного по методу сеток, существуют теоретические оценки. Как правило, эти оценки весьма сложны и применение их затруднительно. Поэтому на практике

№ 1

0	13	49		
20	27	44	73	
24	28	36	44	49
26	27	28	27	13
24	26	24	20	0

№ 2

20	27	46	
26	29	37	46
26	27	29	27
26	26	26	20

№ 3

20	27	46	
26	30	38	46
26	28	30	27
26	26	26	20

№ 4

20	28	47	
26	30	38	47
27	28	30	28
27	27	26	20

№ 5

20	28	47	
27	30	38	47
28	28	30	28
28	28	27	20

№ 6

21	28	47	
27	30	38	47
28	29	30	28
28	28	27	21

№ 7

21	28	47	
27	30	38	47
28	29	30	28
28	28	27	21

№ 7а

0	12	46		
22	28	47	73	
30	33	40	47	46
32	32	33	28	12
32	32	30	22	0

используют двойной пересчет решения с шагами h и $2h$. Если соответствующие результаты совпадают с заданной точностью, то считают, что искомое решение задачи найдено правильно. В противном случае применяют пересчет с шагом $h/2$ и сравнивают полученный результат с прежним результатом, соответствующим шагу h, и т. д. Отдельно следует проанализировать влияние ошибок округления. Схема вычислений должна быть *устойчивой*, т. е. ошибки решения, связанные с округлением, не должны возрастать неограниченно.

§ 8. Понятие о решении задачи Дирихле методом моделирования

Приближенное решение задачи Дирихле может быть получено также с помощью моделирующих устройств. Под *моделированием* с общей точки зрения понимается использование аналогий между физическими явлениями и соответствующими дифференциальными уравнениями.

Задача Дирихле

$$\Delta u(x, y)=0 \text{ при } (x, y)\in G, \quad u(x, y)=\varphi(x, y) \text{ при } (x, y)\in\Gamma,$$

где контур Γ ограничивает область G, приближенно решается с помощью *сеточного электроинтегратора* [16]—[18]. Принцип последнего следующий. В данной области G построим две системы равноотстоящих прямолинейных проводников, параллельных соответственно осям Ox и Oy. Пусть точки пересечения

$$x_i=ih, \quad y_j=jh \quad (i, j=0, \pm1, \pm2, \ldots)$$

(h — расстояние между соседними проводниками) этих проводников являются внутренними узлами сетки, а концы проводников соответствуют граничным узлам. Предположим, что в построенной системе проводников протекает постоянный электрический ток. Тогда значение $u_{ij}=u(x_i, y_j)$ можно интерпретировать как потенциал тока в узле (x_i, y_j) при условии, что в граничных узлах (x_p, y_q) потенциал тока имеет заданную величину $\varphi(x_p, y_q)$. Действительно, в силу закона Кирхгофа суммарное количество электричества, протекающего через каждый внутренний узел (x_i, y_j), равно нулю. Поэтому, обозначая через R омическое сопротивление единицы длины проводника ($R=\text{const}$), на основании закона Ома будем иметь

$$\frac{u(x_i-h, y_j)-u(x_i, y_j)}{Rh}+\frac{u(x_i+h, y_j)-u(x_i, y_j)}{Rh}+$$
$$+\frac{u(x_i, y_j-h)-u(x_i, y_j)}{Rh}+\frac{u(x_i, y_j+h)-u(x_i, y_j)}{Rh}=0.$$

Отсюда

$$u_{ij}=\frac{1}{4}(u_{i-1,j}+u_{i+1,j}+u_{i,j-1}+u_{i,j+1}), \tag{1}$$

если $(x_i, y_j)\in G$, причем

$$u_{pq}=\varphi(x_p, y_q), \text{ если } (x_p, y_q)\in\Gamma. \tag{2}$$

Из формул (1) и (2) вытекает, что u_{ij} являются значениями решения нашей задачи Дирихле, полученного методом сеток по первой основной конечно-разностной схеме (§ 5, (7); § 6, (2)). Величины u_{ij} определяются путем непосредственного измерения потенциалов установившегося тока в узлах (x_i, y_j). Для повышения точности метода применяют специальные приемы ввода граничных условий (см. [16]).

§ 9. Понятие о решении задачи Дирихле методом Монте-Карло

Пусть на плоскости Oxy дана область G с кусочно-гладкой границей Γ. В области G построим квадратную сетку S_h с шагом h:

$$x_i = x_0 + ih, \quad y_j = y_0 + jh \quad (i, j = 0, \pm 1, \pm 2, \ldots), \tag{1}$$

где $\Delta x_i = x_{i+1} - x_i = h$, $\Delta y_j = y_{j+1} - y_j = h$ $(h > 0)$.

Мы предполагаем, что сетка S_h состоит из внутренних узлов и граничных узлов первого рода (см. § 6). Граничные узлы сетки S_h образуют ее *границу* Γ_h. Грубо говоря, граница Γ_h представляет собой линейный ряд точек $M_{pq}(x_p, y_q)$, аппроксимирующий криволинейную границу Γ области G с точностью до h.

Представим себе частицу M, которая совершает равномерное случайное блуждание по узлам сетки (1). А именно, находясь во внутреннем узле $M_{ij}(x_i, y_j)$ сетки S_h, эта частица за один переход с одной и той же вероятностью, равной 1/4, может переместиться в один из четырех соседних узлов: или в $M_{i-1,j}(x_i - h, y_j)$ (шаг влево), или в $M_{i+1,j}(x_i + h, y_j)$ (шаг вправо), или в $M_{i,j-1}(x_i, y_j - h)$ (шаг вниз), или в $M_{i,j+1}(x_i, y_j + h)$ (шаг вверх), причем каждый такой единичный переход совершенно случаен и не зависит от положения частицы и ее прошлой истории. Будем считать, что блуждание частицы M заканчивается, как только эта частица попадает на границу Γ_h; в этом смысле граница Γ_h представляет собой «поглощающий экран». Можно доказать [11], что с вероятностью, равной 1, блуждание точки M через конечное число шагов заканчивается на границе.

Таблица 68

Определение шага частицы в зависимости от выпавшего случайного числа

Случайное число	Характер перемещения
0 или 4	$\Delta x_i = h$ (шаг вправо)
1 » 5	$\Delta y_j = h$ (шаг вверх)
2 » 6	$\Delta x_i = -h$ (шаг влево)
3 » 7	$\Delta y_j = -h$ (шаг вниз)

Если частица M начала свое блуждание с фиксированной внутренней точки $M_{i_0 j_0}$ сетки S_h, то конечная совокупность последовательных положений этой частицы: $M_{i_0 j_0}, M_{i_1 j_1}, \ldots, M_{i_s j_s}$, где $M_{i_k j_k} \overline{\in} \Gamma_h$ $(k = 0, 1, \ldots, s-1)$ и $M_{i_s j_s} \in \Gamma_h$, называется *траекторией частицы* (с s шагами) или *историей блуждания*.

Равномерное случайное блуждание частицы на плоскости можно организовать с помощью равномерно распределенной последовательности одноразрядных случайных чисел [11], [12], принимающих

значения 0, 1, 2, 3, 4, 5, 6, 7, 8, 9. Для этого, например, достаточно производить розыгрыш, т. е. случайную выборку из чисел 0—9, придерживаясь инструкции, указанной в таблице 68; причем числа 8 и 9 переигрываются.

Случайные числа берутся из готовых таблиц или вырабатываются электронной машиной [19]. Последний способ при работе на счетной машине предпочтительнее, так как он позволяет не загружать сильно память машины.

Пусть в точках границы Γ области G определена некоторая функция $\varphi(x, y)$. Перенесем эти значения на границу Γ_h сетки S_h. Например (см. § 6), для каждого граничного узла $M_{pq}(x_p, y_q) \in \Gamma_h$ определим ближайшую по горизонтали (или вертикали) точку $N \in \Gamma$ и положим

$$\varphi(M_{pq}) = \varphi(N).$$

Для краткости введем обозначение

$$\varphi_{pq} = \varphi(x_p, y_q).$$

Пусть $P(i, j; p, q)$ — вероятность того, что траектория частицы, вышедшей из узла M_{ij} сетки S_h, закончится в граничном узле $M_{pq} \in \Gamma_h$. Так как блуждание точки неизбежно заканчивается на границе Γ_h в первой же точке выхода ее на границу, то

$$\sum_{p,q} P(i, j; p, q) = 1, \tag{2}$$

где суммирование распространяется на все точки M_{pq} границы Γ_h, причем

$$P(p', q'; p, q) = \begin{cases} 1 \text{ при } p' = p,\ q' = q, \\ 0 \quad » \quad |p' - p| + |q' - q| \neq 0, \end{cases} \tag{3}$$

где $M_{p'q'}$ — граничный узел.

Составим сумму

$$v_{ij} = \sum_{p,q} P(i, j; p, q)\, \varphi_{pq}, \tag{4}$$

где точка $M_{pq}(x_p, y_q)$ пробегает всю границу Γ_h. Если функцию $\varphi(x, y)$ рассматривать как случайную величину, принимающую значения φ_{pq} на границе Γ_h, то сумма (4) представляет собой математическое ожидание (среднее значение) функции $\varphi(x, y)$ на границе Γ_h для траекторий, начинающихся в точке $M_{ij}(x_i, y_j)$ («премия за выход на границу» из начальной точки M_{ij}).

Частица, начавшая свое случайное блуждание из внутреннего узла M_{ij}, после первого шага с вероятностью, равной 1/4, попадает в один из четырех соседних узлов. Поэтому случайные блуждания, начинающиеся в узле M_{ij}, в зависимости от вида траекторий распадаются на четыре категории новых случайных блужданий:

I. M_{ij}, $M_{i-1,j}$, ... III. M_{ij}, $M_{i,j-1}$, ...

II. M_{ij}, $M_{i+1,j}$, ... IV. M_{ij}, $M_{i,j+1}$, ...

По формуле полной вероятности имеем

$$P(i, j; p, q) = \frac{1}{4} P(i-1, j; p, q) + \frac{1}{4} P(i+1, j; p, q) + \\ + \frac{1}{4} P(i, j-1; p, q) + \frac{1}{4} P(i, j+1; p, q). \quad (5)$$

Отсюда, умножая обе части равенства (5) на граничные значения φ_{pq} и суммируя по всем возможным значениям p и q, на основании формулы (4) получим (см. [11])

$$v_{ij} = \frac{1}{4} (v_{i-1, j} + v_{i+1, j} + v_{i, j-1} + v_{i, j+1}). \quad (6)$$

Кроме того, в силу формулы (3) имеем

$$v_{pq} = \varphi_{pq}, \quad (7)$$

если точка $M_{pq} \in \Gamma_h$.

Рассмотрим теперь задачу Дирихле об отыскании функции $u = u(x, y)$, гармонической области G и принимающей на ее границе Γ заданные непрерывные значения $\varphi(x, y)$. Согласно методу сеток (§ 6) эта задача сводится к нахождению значений $u_{ij} = u(x_i, y_j)$ искомой функции $u(x, y)$ во внутренних узлах $M_{ij}(x_i, y_j)$ некоторой сетки S_h при условии, что значения в граничных узлах $M_{pq}(x_p, y_q) \in \Gamma_h$ известны и равны $\varphi_{pq} = \varphi(x_p, y_q)$. Неизвестные u_{ij} определяются из системы линейных уравнений (см. § 6, формула (2))

$$\left.\begin{aligned} u_{ij} &= \frac{1}{4} (u_{i-1, j} + u_{i+1, j} + u_{i, j-1} + u_{i, j+1}), \\ u_{pq} &= \varphi_{pq}. \end{aligned}\right\} \quad (8)$$

Сравнивая формулы (8) с формулами (6), (7), мы усматриваем, что они совпадают с точностью до обозначений. Следовательно, искомые неизвестные u_{ij} можно рассматривать как математические ожидания v_{ij}.

Величины v_{ij} допускают экспериментальное определение. Рассмотрим достаточно большое число N равномерных случайных блужданий частицы по узлам сетки S_h, исходящих из фиксированного узла M_{ij} и заканчивающихся на границе Γ_h. Пусть $(x_p^{(k)}, y_q^{(k)})$ $(k = 1, 2, \ldots, N)$ — соответствующие точки выхода частицы на границу Γ_h. Заменяя математическое ожидание v_{ij} эмпирическим математическим ожиданием, будем иметь

$$u_{ij} = v_{ij} \approx \frac{1}{N} \sum_{k=1}^{N} \varphi(x_p^{(k)}, y_q^{(k)}). \quad (9)$$

Формула (9) дает статистическую оценку величины u_{ij} и может быть применена для приближенного решения задачи Дирихле. Метод решения задач, основанный на использовании случайных величин, получил общее название *метода Монте-Карло* (см. [13]).

Заметим, что с помощью формулы (9) можно непосредственно найти приближенное значение u_{ij} решения задачи Дирихле в единственной фиксированной точке M_{ij} сетки S_h, не зная решения задачи для остальных точек сетки. Этим обстоятельством метод Монте-Карло для задачи Дирихле резко отличается от обычных стандартных способов решения этой задачи.

Интересно отметить, что вероятность $P(i, j; p, q)$, в силу формулы (4), представляет собой аналог *функции Грина* для задачи Дирихле в области S_h. Эта величина может быть найдена экспериментально на основании формулы (9), если задать следующие граничные условия:

$$\varphi(x_p, y_q) = 1, \qquad \varphi(x_{p'}, y_{q'}) = 0$$

при $|p'-p|+|q'-q| \neq 0$.

Построив такую функцию Грина, мы получаем возможность, применяя формулу (9), просто находить приближенное решение задачи Дирихле для области G с данной границей Γ при любых граничных значениях $\varphi(x, y)$.

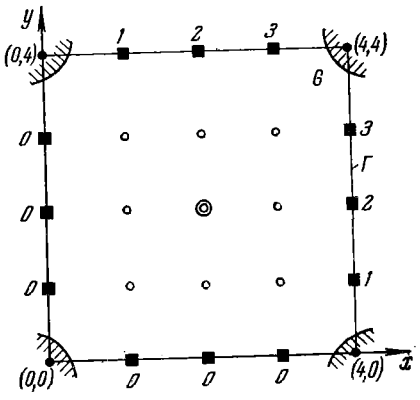

Рис. 71.

Недостатком рассмотренного варианта метода Монте-Карло для задачи Дирихле является слабая сходимость по вероятности при $N \to \infty$ эмпирического математического ожидания

$$\tilde{v}_{ij} = \frac{1}{N} \sum_{k=1}^{N} \varphi(x_p^{(k)}, y_q^{(k)})$$

к математическому ожиданию v_{ij}. Чтобы устранить это неблагоприятное обстоятельство, используют различные модификации случайных блужданий (см. [11]). Кроме того, при решении задачи полезно учитывать также, что блуждание частицы M, начинающееся в точке M_{ij}, автоматически является случайным блужданием частицы, начинающимся в любой промежуточной точке траектории этой частицы.

Пример 1. Методом Монте-Карло найти $u(2, 2)$, если

$$\Delta u(x, y) = 0 \text{ в области } G\{0 \leqslant x \leqslant 4;\ 0 \leqslant y \leqslant 4\} \tag{10}$$

и

$$\begin{aligned} &u(x, 0) = 0, \quad 0 \leqslant x \leqslant 4; \qquad u(x, 4) = x, \quad 0 \leqslant x \leqslant 4; \\ &u(4, y) = y, \quad 0 \leqslant y \leqslant 4; \qquad u(0, y) = 0, \quad 0 \leqslant y \leqslant 4. \end{aligned} \tag{11}$$

Решение. Для квадрата G с границей Γ построим квадратную сетку S_h с шагом $h=1$ (рис. 71). Рассматриваем серию равномерных блужданий частицы по узлам нашей сетки S_h, исходящих из начального положения (2, 2)

(на рис. 71 эта точка отмечена двойным кружком) и заканчивающихся на границе Г в граничных узлах сетки G_h (на рисунке эти узлы отмечены квадратиками). Для организации случайных блужданий используем равномерно распределенную последовательность случайных чисел [14] (см.таблицу 69).

Перемещение частицы определялось согласно указанной выше инструкции (таблица 68), причем выпадение чисел 8 и 9 рассматривалось как стояние частицы на месте. Например, последовательно разыгрывая числа таблицы 69,

Таблица 69

Последовательность случайных чисел

57705	71618	73710	70131	16961	53324	43166
26275	05926	66289	35483	09393	30304	55186
64003	20514	00188	55709	86977	31303	11578
93045	93011	42844	52906	09461	99602	69962

начиная с первого, имеем:

случайное число 5 — шаг вверх (2, 2) $\longrightarrow$ (2, 3);
случайное число 7 — шаг вниз (2, 3) $\longrightarrow$ (2, 2);
случайное число 7 — шаг вниз (2, 2) $\longrightarrow$ (2, 1);
случайное число 0 — шаг вправо (2, 1) $\longrightarrow$ (3, 1);
случайное число 5 — шаг вверх (3, 1) $\longrightarrow$ (3, 2);
случайное число 7 — шаг вниз (3, 2) $\longrightarrow$ (3, 1);
случайное число 1 — шаг вверх (3, 1) $\longrightarrow$ (3, 2);
случайное число 6 — шаг влево (3, 2) $\longrightarrow$ (2, 2);
случайное число 1 — шаг вверх (2, 2) $\longrightarrow$ (2, 3);
случайное число 8 — на месте (2, 3) $\longrightarrow$ (2, 3);
случайное число 7 — шаг вниз (2, 3) $\longrightarrow$ (2, 2);
случайное число 3 — шаг вниз (2, 2) $\longrightarrow$ (2, 1);
случайное число 7 — шаг вниз (2, 1) $\longrightarrow$ (2, 0).

В узле (2, 0) блуждание прекращается, так как частица вышла на границу. Далее переходим ко второму блужданию частицы, снова отправляясь из той же начальной точки (2, 2), и т. д. Траектории $N=20$ случайных блужданий приведены в таблице 70.

На основании формулы (9) имеем

$$u(2,\ 2) \approx \frac{1}{20} \cdot \sum = \frac{1}{20} \cdot 20 = 1.$$

Заметим, что в данном случае для задачи Дирихле известно точное решение

$$u(x,\ y) = \frac{xy}{4}.$$

Поэтому $u(2,\ 2) = \frac{2 \cdot 2}{4} = 1$. Таким образом, мы случайно получили точное значение $u(2,\ 2)$.

Таблица 70

Решение задачи Дирихле (10)—(11) методом Монте-Карло

№ блуждания	Траектория блуждания	Значение функции $u(x, y)$ в точке выхода на границу
1	(2,2) → (2,3) → (2,2) → (2,1) → → (3,1) → (3,2) → (3,1) → (3,2) → → (2,2) → (2,3) → (2,3) → (2,2) → → (2,1) → (2,0);	0
2	(2,2) → (2,3) → (3,3) → (3,2) → → (4,2);	2
3	(2,2) → (2,3) → (2,2) → (2,3) → → (2,4);	2
4	(2,2) → (1,2) → (1,2) → (0,2).	0
5	(2,2) → (2,3) → (2,4);	2
6	(2,2) → (2,1) → (2,0);	0
7	(2,2) → (1,2) → (2,2) → (3,2) → → (3,1) → (3,2) → (2,2) → (1,2) → → (0,2);	0
8	(2,2) → (1,2) → (0,2);	0
9	(2,2) → (2,1) → (2,2) → (3,2) → → (3,3) → (3,3) → (2,3) → (1,3) → → (0,3);	0
10	(2,2) → (1,2) → (0,2);	0
11	(2,2) → (2,2) → (2,2) → (2,1) → → (2,2) → (3,2) → (3,2) → (3,1) → → (4,1);	1
12	(2,2) → (2,2) → (2,1) → (2,1) → → (2,0);	0
13	(2,2) → (2,1) → (3,1) → (3,0);	0
14	(2,2) → (3,2) → (4,2);	2
15	(2,2) → (2,3) → (2,4);	2
16	(2,2) → (2,3) → (2,3) → (1,3) → → (0,3);	0
17	(2,2) → (3,2) → (4,2);	2
18	(2,2) → (3,2) → (3,1) → (2,1) → → (3,1) → (3,2) → (3,3) → (4,3);	3
19	(2,2) → (3,2) → (4,2);	2
20	(2,2) → (2,3) → (2,3) → (2,3) → → (2,4);	2
	Σ	20

§ 10. Метод сеток для уравнения параболического типа

В качестве примера уравнения параболического типа остановимся на уравнении теплопроводности для однородного стержня $0 \leqslant x \leqslant l$:

$$\frac{\partial u}{\partial t} = a^2 \frac{\partial^2 u}{\partial x^2}, \tag{1}$$

где $u = u(x, t)$ — температура и t — время (см. § 1, пример 3). В дальнейшем для простоты будем полагать $a = 1$ (к такому случаю всегда можно прийти путем введения нового времени $\tau = a^2 t$).

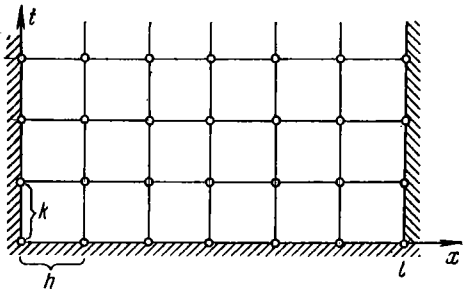

Рис. 72.

Итак, рассмотрим уравнение

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}. \tag{2}$$

Пусть, кроме того, в начальный момент времени $t = 0$ задано распределение температуры $u(x, 0) = f(x)$ и законы изменения температуры в зависимости от времени (тепловые режимы) на концах стержня $x = 0$ и $x = l$:

$$u(0, t) = \varphi(t), \qquad u(l, t) = \psi(t).$$

Требуется найти распределение температуры $u = u(x, t)$ вдоль стержня в любой момент времени t. Решим эту смешанную задачу методом сеток [6], [7]. Для этого рассмотрим пространственно-временную систему координат $\{x, t\}$ (рис. 72). В полуполосе $t \geqslant 0$, $0 \leqslant x \leqslant l$ построим прямоугольную сетку

$$x = ih \quad (i = 0, 1, 2, \ldots, n), \qquad t = jk \quad (j = 0, 1, 2, \ldots),$$

где $h = \frac{l}{n}$ (n — целое) — шаг вдоль оси Ox и $k = \sigma h^2$ (σ — постоянная) — шаг вдоль оси Ot, вообще говоря, различны. Величина σ будет выбрана ниже. Введя обозначения

$$x_i = ih; \quad t_j = jk; \quad u_{ij} = u(x_i, t_j)$$

и заменяя уравнение (2) конечно-разностным уравнением, будем иметь

$$\frac{u_{i,\,j+1} - u_{ij}}{\sigma h^2} = \frac{u_{i+1,\,j} - 2u_{ij} + u_{i-1,\,j}}{h^2}. \tag{3}$$

Отсюда

$$u_{i,\,j+1} = \sigma u_{i-1,\,j} + (1 - 2\sigma) u_{ij} + \sigma u_{i+1\,j}. \tag{4}$$

Из рассмотрения формулы (4) ясно, что, зная значения функции

$u(x, t)$ в точках j-го слоя $t = jk$, с помощью этой формулы можно вычислить значения $u(x, t)$ в точках следующего $(j+1)$-го слоя $t = (j+1)k$ (рис. 73). При вычислении пользуются четырьмя соседними узлами — *явная схема вида* $\overset{*}{***}$ (схема 1).

Таким образом, исходя из начального слоя $t = 0$, значения $u(x, t)$ для которого определяются из начального условия

$$u(x_i, 0) = f(x_i) \qquad (i = 0, 1, 2, \ldots, n),$$

и используя значения функции $u(x, t)$ в крайних узлах $(0, t_j)$, (l, t_j) $(j = 0, 1, \ldots)$, определяемые граничными условиями

$$u(0, t_j) = \varphi(t_j),$$
$$u(l, t_j) = \psi(t_j),$$

по формуле (4) последовательно вычисляем:

$$u(x_i, t_1);\ u(x_i, t_2);$$
$$u(x_i, t_3);\ \ldots$$
$$(i = 0, 1, \ldots, n),$$

Схема 1

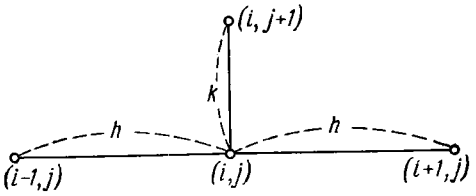

Рис. 73.

т. е. находим значения искомой функции $u(x, t)$ во всех узлах полуполосы.

Остается разумно выбрать величину σ. При этом будем исходить из требования, чтобы ошибка при замене дифференциального уравнения (2) конечно-разностным уравнением (3) была наименьшей.

Введем обозначения:

$$L[u] = \frac{\partial^2 u}{\partial x^2} - \frac{\partial u}{\partial t},$$

$$L_h[u] = \frac{1}{h^2}\left[(u_{i+1,j} - 2u_{ij} + u_{i-1,j}) - \frac{1}{\sigma}(u_{i,j+1} - u_{ij})\right],$$

где $L_h(u)$ — *конечно-разностный оператор*, соответствующий дифференциальному оператору $L[u]$.

Разность

$$R_h[u] = L_h[u] - L[u],$$

называемая *ошибкой аппроксимации*, есть та погрешность, которая получается при замене оператора $L[u]$ оператором $L_h[u]$. Вычислим эту погрешность в узлах (x_i, t_j) сетки для функции $u(x, y)$, являющейся решением уравнения (2). При этом $L[u] = 0$ и

$$R_h[u] = L_h[u]. \tag{5}$$

Учитывая, что

$$u_{i+1,j} = u(x_i + h, t_j), \qquad u_{i-1,j} = u(x_i - h, t_j),$$
$$u_{i,j+1} = u(x_i, t_j + \sigma h^2),$$

и разлагая $L_h[u]$ по формуле Тейлора в окрестности точки (x_i, t_j), ограничиваясь членами порядка h^6, находим

$$L_h[u]=\frac{1}{h^2}\left\{\left(u_{ij}+h\frac{\partial u_{ij}}{\partial x}+\frac{h^2}{2!}\frac{\partial^2 u_{ij}}{\partial x^2}+\frac{h^3}{3!}\frac{\partial^3 u_{ij}}{\partial x^3}+\frac{h^4}{4!}\frac{\partial^4 u_{ij}}{\partial x^4}+\right.\right.$$
$$+\frac{h^5}{5!}\frac{\partial^5 u_{ij}}{\partial x^5}+\frac{h^6}{6!}\frac{\partial^6 u_{ij}}{\partial x^6}-2u_{ij}+u_{ij}-h\frac{\partial u_{ij}}{\partial x}+\frac{h^2}{2!}\frac{\partial^2 u_{ij}}{\partial x^2}-$$
$$\left.-\frac{h^3}{3!}\frac{\partial^3 u_{ij}}{\partial x^3}+\frac{h^4}{4!}\frac{\partial^4 u_{ij}}{\partial x^4}-\frac{h^5}{5!}\frac{\partial^5 u_{ij}}{\partial x^5}+\frac{h^6}{6!}\frac{\partial^6 u_{ij}}{\partial x^6}\right)-$$
$$\left.-\frac{1}{\sigma}\left[u_{ij}+\frac{\partial u_{ij}}{\partial t}\sigma h^2+\frac{(\sigma h^2)^2}{2!}\frac{\partial^2 u_{ij}}{\partial t^2}+\frac{(\sigma h^2)^3}{3!}\frac{\partial^3 u_{ij}}{\partial t^3}-u_{ij}\right]\right\}+O(h^6).$$

Отсюда после приведения подобных членов будем иметь

$$L_h[u]=\left(\frac{\partial^2 u_{ij}}{\partial x^2}-\frac{\partial u_{ij}}{\partial t}\right)+h^2\left(\frac{1}{12}\frac{\partial^4 u_{ij}}{\partial x^4}-\frac{\sigma}{2}\frac{\partial^2 u_{ij}}{\partial t^2}\right)+$$
$$+h^4\left(\frac{1}{360}\frac{\partial^6 u_{ij}}{\partial x^6}-\frac{\sigma^2}{6}\frac{\partial^3 u_{ij}}{\partial t^3}\right)+O(h^6). \qquad (6)$$

Так как $u(x, t)$ есть решение уравнения (2), то

$$\frac{\partial^2 u_{ij}}{\partial x^2}=\frac{\partial u_{ij}}{\partial t};\quad \frac{\partial^4 u_{ij}}{\partial x^4}=\frac{\partial^2 u_{ij}}{\partial t^2};\quad \frac{\partial^6 u_{ij}}{\partial x^6}=\frac{\partial^3 u_{ij}}{\partial t^3}.$$

Заменяя в (6) частные производные по t равными им частными производными по x, получаем

$$L_h[u]=h^2\left(\frac{1}{12}-\frac{\sigma}{2}\right)\frac{\partial^4 u_{ij}}{\partial x^4}+h^4\left(\frac{1}{360}-\frac{\sigma^2}{6}\right)\frac{\partial^6 u_{ij}}{\partial x^6}+O(h^6). \qquad (7)$$

Выберем число σ так, чтобы первая скобка формулы (7) обратилась в нуль, т. е. положим $\sigma/2=1/12$ и, следовательно, $\sigma=1/6$. При этом значении σ будем иметь

$$L_h[u]=h^4\left(\frac{1}{360}-\frac{1}{216}\right)\frac{\partial^6 u_{ij}}{\partial x^6}+O(h^6)=-\frac{h^4}{540}\cdot\frac{\partial^6 u_{ij}}{\partial x^6}+O(h^6).$$

В силу (5) выполнено равенство $R_h[u]=L_h[u]$. Поэтому при таком выборе σ для погрешности $R_h[u]$ получаем оценку $R_h[u]=O(h^4)$, тогда как при другом выборе числа σ имеем $R_h[u]=O(h^2)$. В этом смысле значение $\sigma=1/6$ является для расчетной схемы I наилучшим.

Соответствующая расчетная формула (4) при таком выборе σ окончательно принимает вид

$$u_{i,j+1}=\frac{1}{6}(u_{i-1,j}+4u_{ij}+u_{i+1,j}). \qquad (8)$$

Отметим, что оценка ошибки аппроксимации $R_h[u]$ в общем случае для граничных узлов (x_i, t_j) не годится.

Пример 1. Найти решение уравнения теплопроводности

$$\frac{\partial u}{\partial t}=\frac{\partial^2 u}{\partial x^2}$$

при следующих начальных и краевых условиях [7]:

$$u(x,\ 0)=4x(1-x) \quad (0\leqslant x\leqslant 1);$$
$$u(0,\ t)=0 \quad \text{и} \quad u(1,\ t)=0 \quad (0\leqslant t<\infty).$$

Решение. Заметим, что начальные и краевые условия задачи симметричны относительно прямой $x=\frac{1}{2}$. Поэтому и решение $u(x,\ t)$ будет симметрично относительно этой прямой.

Для расчетов полагаем $h=1/10$ и, следовательно, $k=\sigma h^2=1/600$ и строим систему узлов $(x_i,\ t_j)$, где $x_i=0{,}1i$, $t_j=j/600$. Результаты вычислений удобно расположить в прямоугольной таблице 71.

Таблица 71

Решение уравнения теплопроводности методом сеток

j \ i	$i=0$	$i=1$	$i=2$	$i=3$	$i=4$	$i=5$	$i=6$	$i=7$	$i=8$	$i=9$	$i=10$
$j=0$	0	0,360	0,040	0,840	0,960	1,000	0,960	0,840	0,640	0,360	0
$j=1$	0	0,347	0,627	0,827	0,947	0,987	0,947	0,827	0,627	0,347	0
$j=2$	0	0,336	0,613	0,813	0,933	0,973	0,933	0,813	0,613	0,336	0
$j=3$	0	0,326	0,600	0,800	0,920	0,960	0,920	0,800	0,600	0,326	0
$j=4$	0	0,317	0,588	0,787	0,907	0,947	0,907	0,787	0,588	0,317	0
$j=5$	0	0,309	0,576	0,774	0,894	0,934	0,894	0,774	0,576	0,309	0
$j=6$	0	0,302	0,564	0,761	0,881	0,921	0,881	0,761	0,564	0,302	0

Начальная строка этой таблицы ($j=0$) заполняется на основании заданных начальных условий

$$u(x_i,\ 0)=4x_i(1-x_i)=0{,}4i(1-0{,}1i) \quad (i=0,\ 1,\ 2,\ \ldots,\ 10).$$

В первый ($i=0$) и последний ($i=10$) столбцы вписываются данные граничных условий

$$u(0,\ t_j)=u(1,\ t_j)=0 \quad (j=0,\ 1,\ 2,\ \ldots).$$

Остальные строки $j=1,\ 2,\ 3,\ \ldots$ таблицы последовательно заполняются с помощью применения расчетной формулы (8). При этом, конечно, следует учитывать симметрию искомой функции u.

§ 11. Устойчивость конечно-разностной схемы для решения уравнения теплопроводности

При использовании конечно-разностной схемы для решения краевой задачи возникает важный вопрос об *устойчивости* такой схемы. Под этим понимается следующее: конечно-разностная схема называется *устойчивой*, если малые погрешности, допущенные в процессе решения, затухают или во всяком случае остаются малыми при неограни-

ченном увеличении номера текущего слоя (строгое определение понятия устойчивости см. в [8]). В противном случае схема называется *неустойчивой*. Ясно, что неустойчивая конечно-разностная схема противопоказана для вычислений, так как неизбежные незначительные ошибки, например погрешности округлений, могут создать большие отклонения от точного решения краевой задачи и привести к результатам, не имеющим ничего общего с действительностью.

Выясним условия устойчивости приведенной выше конечно-разностной схемы (§ 10, (4)) для уравнения теплопроводности

$$\frac{\partial u}{\partial t}=\frac{\partial^2 u}{\partial x^2} \tag{1}$$

с заданными граничными и начальными условиями (*смешанная краевая задача*):

$$\left.\begin{aligned} u(x,\ 0)&=f(x),\\ u(0,\ t)&=\varphi(t),\\ u(l,\ t)&=\psi(t). \end{aligned}\right\} \tag{2}$$

Пусть $\Delta x_i=h$ и $\Delta t_j=k$, где $h=l/n$, $k=\sigma h^2$. Переходя к конечным разностям в уравнении (1) (см. § 10), будем иметь

$$\sigma u(x_i+h,\ t_j)+(1-2\sigma)\,u(x_i,\ t_j)+\sigma u(x_i-h,\ t_j)-\\ -u(x_i,\ t_j+k)=0. \tag{3}$$

В граничных узлах сетки $x_i,\ t_j\in\Gamma$ выполнены следующие условия:

$$u(x_i,\ 0)=f(x_i),\qquad u(0,\ t_j)=\varphi(t_j),\qquad u(l,\ t_j)=\psi(t_j). \tag{4}$$

Предположим, что в точках начального слоя $t=0$ допущена ошибка $\varepsilon_{i,0}$, т. е. $u(x_i,\ 0)=f(x_i)+\varepsilon_{i,0}$, и пусть $v(x_i,\ t_j)$ — решение уравнения (3):

$$\sigma v(x_i+h,\ t_j)+(1-2\sigma)\,v(x_i,\ t_j)+\sigma v(x_i-h,\ t_j)-\\ -v(x_i,\ t_j+k)=0, \tag{5}$$

удовлетворяющее граничным условиям, содержащим ошибку:

$$v(x_i,\ 0)=f(x_i)+\varepsilon_{i,0},\qquad v(0,\ t_j)=\varphi(t_j),\qquad v(l,\ t_j)=\psi(t_j). \tag{6}$$

Нас интересует, как изменяется погрешность

$$w(x_i,\ t_j)=v(x_i,\ t_j)-u(x_i,\ t_j)$$

при неограниченном возрастании номера j. Вычитая из уравнения (5) уравнение (3), для погрешности $w(x_i,\ t_j)$ получим конечно-разностное уравнение

$$\sigma w(x_i+h,\ t_j)+(1-2\sigma)\,w(x_i,\ t_j)+\sigma w(x_i-h,\ t_j)-\\ -w(x_i,\ t_j+k)=0. \tag{7}$$

На границе Г области имеем

$$\left.\begin{aligned} &w(x_i,\ 0)=v(x_i,\ 0)-u(x_i,\ 0)=\varepsilon_{i,0},\\ &w(0,\ t_j)=v(0,\ t_j)-u(0,\ t_j)=0,\\ &w(l,\ t_j)=v(l,\ t_j)-u(l,\ t_j)=0. \end{aligned}\right\} \tag{8}$$

Частное решение уравнения (7) будем искать в виде

$$w(x_i,\ t_j)=\lambda^{t_j}\sin px_i, \tag{9}$$

где числа λ и p $(p>0)$ подберем так, чтобы выражение (9) удовлетворяло уравнению (7) и однородным краевым условиям

$$w(0,\ t_j)=0, \qquad w(l,\ t_j)=0. \tag{10}$$

На основании этих условий имеем

$$\lambda^{t_j}\sin p0=0, \qquad \lambda^{t_j}\sin pl=0,$$

откуда вытекает, что

$$pl=m\pi \quad \text{и} \quad p=\frac{m\pi}{l} \quad (m=1,\ 2,\ \ldots).$$

Следовательно,

$$w(x_i,\ t_j)=\lambda^{t_j}\sin\frac{m\pi x_i}{l}.$$

Подставляя это выражение в уравнение (7), будем иметь

$$\sigma\lambda^{t_j}\sin\frac{m\pi}{l}(x_i+h)+(1-2\sigma)\lambda^{t_j}\sin\frac{m\pi}{l}x_i+$$
$$+\sigma\lambda^{t_j}\sin\frac{m\pi}{l}(x_i-h)-\lambda^{t_j+k}\sin\frac{m\pi}{l}x_i=0$$

или

$$\sigma\left[\sin\frac{m\pi}{l}(x_i+h)-2\sin\frac{m\pi}{l}x_i+\sin\frac{m\pi}{l}(x_i-h)\right]+$$
$$+(1-\lambda^k)\sin\frac{m\pi}{l}x_i=0. \tag{11}$$

Выражение, стоящее в квадратных скобках равенства (11), приведем к виду, удобному для логарифмирования. Имеем

$$\sin\frac{m\pi}{l}(x_i+h)-2\sin\frac{m\pi}{l}x_i+\sin\frac{m\pi}{l}(x_i-h)=2\sin\frac{m\pi}{l}x_i\cos\frac{m\pi}{l}h-$$
$$-2\sin\frac{m\pi}{l}x_i=2\sin\frac{m\pi}{l}x_i\left(\cos\frac{m\pi}{l}h-1\right)=-4\sin\frac{m\pi}{l}x_i\sin^2\frac{m\pi}{2l}h.$$

Подставляя этот результат в равенство (11) и сокращая обе части его на общий множитель $\sin\frac{m\pi}{l}x_i$, получим

$$-4\sigma\sin^2\frac{m\pi}{2l}h+(1-\lambda^k)=0.$$

Отсюда

$$\lambda^k = 1 - 4\sigma \sin^2 \frac{m\pi}{2l} h$$

и

$$\lambda = \left(1 - 4\sigma \sin^2 \frac{m\pi}{2l} h\right)^{\frac{1}{k}}. \tag{12}$$

Заметим, что λ не зависит от точки (x_i, t_j). Таким образом, для однородного уравнения (7) получаем линейно независимые решения вида

$$w_m(x_i, t_j) = \left(1 - 4\sigma \sin^2 \frac{m\pi}{2l} h\right)^{\frac{t_j}{k}} \sin \frac{m\pi}{l} x_i \qquad (m = 1, 2, \ldots, n-1),$$

причем каждое решение удовлетворяет однородным краевым условиям (10).

Линейная комбинация этих решений

$$w(x_i, t_j) = \sum_{m=1}^{n-1} c_m w_m(x_i, t_j) = \\ = \sum_{m=1}^{n-1} c_m \left(1 - 4\sigma \sin^2 \frac{m\pi}{2l} h\right)^{\frac{t_j}{k}} \sin \frac{m\pi}{l} x_i \tag{13}$$

также является решением уравнения (7), удовлетворяющим при любых значениях коэффициентов c_m условиям (10). Эти коэффициенты подбираются так, чтобы выполнялось первое условие (8), т. е. чтобы

$$w(x_i, 0) = \varepsilon_{i,0} \qquad (i = 1, 2, \ldots, n-1).$$

Для устойчивости рассматриваемой конечно-разностной схемы (3) необходимо, чтобы при любых значениях постоянных $c_1, c_2, \ldots, c_{n-1}$ функция $w(x_i, t_j)$, определяемая равенством (13), оставалась ограниченной при $t_j \longrightarrow \infty$. Для этого достаточно, чтобы для всех m было выполнено неравенство

$$\left|1 - 4\sigma \sin^2 \frac{m\pi h}{2l}\right| \leqslant 1.$$

Отсюда

$$-1 \leqslant 1 - 4\sigma \sin^2 \frac{m\pi h}{2l} \leqslant 1$$

и

$$0 \leqslant \sigma \sin^2 \frac{m\pi}{2l} \leqslant \frac{1}{2} \qquad (m = 1, 2, \ldots, n-1). \tag{14}$$

Неравенство (14) заведомо будет выполнено, если величина σ удовлетворяет условию

$$0 < \sigma \leqslant \frac{1}{2}. \tag{15}$$

Полученные неравенства дают достаточные условия устойчивости рассмотренной ранее конечно-разностной схемы (3) для решения смешанной задачи в случае уравнения теплопроводности. Заметим, что выбранное в § 10 значение $\sigma = 1/6$ удовлетворяет неравенству (15).

Замечание. Взяв $\sigma = 1/2$, мы получаем очень удобное конечно-разностное уравнение

$$2u_{i,j+1} = u_{i-1,j} + u_{i+1,j}. \tag{16}$$

Схема устойчива, однако в этом случае порядок отклонения уравнения (16) от уравнения теплопроводности равен $O(h^2)$, в то время как при $\sigma = 1/6$ этот порядок равен $O(h^4)$.

§ 12. Метод прогонки для уравнения теплопроводности

В предыдущих параграфах мы видели, что для устойчивости конечно-разностной схемы для уравнения теплопроводности шаги $h = \Delta x_i$ и $k = \Delta t_j$ должны быть неодинаковы, причем выбор шага h для пространственной координаты x накладывает определенные ограничения на величину шага k для временной координаты t. Важность этого обстоятельства была отмечена Курантом, Фридрихсом и Леви. Так как при устойчивой схеме шаг k имеет порядок $O(h^2)$, причем отношение $\sigma = k/h^2$ ограничено сверху, то при малом h продвижение решения $u(x, t)$ по t весьма незначительно и объем работы чрезвычайно велик. Например, приняв $h = 0{,}1$ и полагая $k = \sigma h^2 = 1/600$, получим, что для описания процесса распространения тепла за единичный промежуток времени $0 \leqslant t \leqslant 1$ требуется таблица, содержащая 600 строк!

Схема II

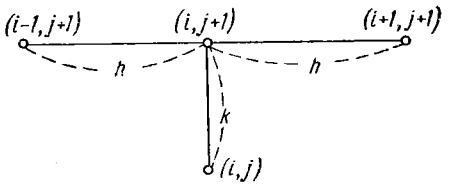

Рис. 74.

Мы сейчас укажем другую устойчивую вычислительную схему, для которой отношение k/h^2 не является ограниченным сверху и поэтому шаг $k = \Delta t_j$ временной координаты может быть выбран сравнительно крупным.

Рассмотрим по-прежнему в области $G\{0 \leqslant x \leqslant l,\ 0 \leqslant t < +\infty\}$ приведенное уравнение теплопроводности

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} \tag{1}$$

с граничными и начальными условиями

$$u(0, t) = \varphi(t), \quad u(l, t) = \psi(t), \quad u(x, 0) = f(x). \tag{2}$$

Построим в области G прямоугольную сетку

$$x_i = ih, \quad t_j = jk \quad (i = 0, 1, 2, \ldots, n; \quad j = 0, 1, 2, \ldots),$$

где $h = l/n$ (n — целое) и k — некоторая положительная величина. Пусть $u_{ij} = u(x_i, t_j)$.

Используя приближенную симметричную формулу для второй производной по x и применяя формулу численного дифференцирования по t «назад», для $(j+1)$-го слоя сетки вместо дифференциального уравнения (1) будем иметь следующее конечно-разностное уравнение [10]:

$$\frac{u_{i-1,\,j+1} - 2u_{i,\,j+1} + u_{i+1,\,j+1}}{h^2} = \frac{u_{i,\,j+1} - u_{i,\,j}}{k}$$

$$(i = 1, 2, \ldots, n-1; \quad j = 0, 1, 2, \ldots),$$

или

$$u_{i-1,\,j+1} - (2+s)\,u_{i,\,j+1} + u_{i+1,\,j+1} = -su_{ij}, \tag{3}$$

где $s = h^2/k$.

Таким образом, здесь используется схема II вида *** (*неявная схема*) (рис. 74).

Из граничных условий (2) получаем

$$u_{0,\,j+1} = \varphi(t_{j+1}), \qquad u_{n,\,j+1} = \psi(t_{j+1}). \tag{4}$$

Систему (3)—(4) будем решать *методом прогонки*. Пусть

$$u_{i,\,j+1} = a_{i,\,j+1}(b_{i,\,j+1} + u_{i+1,\,j+1}) \tag{5}$$

и, следовательно,

$$u_{i-1,\,j+1} = a_{i-1,\,j+1}(b_{i-1,\,j+1} + u_{i,\,j+1}). \tag{6}$$

Подставляя выражение (6) в формулу (3), будем иметь

$$a_{i-1,\,j+1}(b_{i-1,\,j+1} + u_{i,\,j+1}) - (2+s)\,u_{i,\,j+1} + u_{i+1,\,j+1} = -su_{ij},$$

отсюда

$$u_{i,\,j+1} = \frac{a_{i-1,j+1}\,b_{i-1,\,j+1} + su_{ij} + u_{i+1,\,j+1}}{2+s-a_{i-1,\,j+1}}.$$

Сравнивая это выражение с формулой (5), получим

$$a_{i,\,j+1} = \frac{1}{2+s-a_{i-1,\,j+1}}, \quad b_{i,\,j+1} = a_{i-1,\,j+1}b_{i-1,\,j+1} + su_{ij} \tag{7}$$

$$(i = 2, 3, \ldots, n).$$

При $i = 1$ из формул (3) и (5) имеем

$$u_{0,\,j+1} - (2+s)\,u_{1,\,j+1} + u_{2,\,j+1} = -su_{1j}$$

и

$$u_{1,\,j+1} = a_{1,\,j+1}(b_{1,\,j+1} + u_{2,\,j+1}). \tag{8}$$

Отсюда, используя граничные условия, получаем

$$u_{1,\,j+1}=\frac{\varphi(t_{j+1})+su_{1j}+u_{2,\,j+1}}{2+s}. \tag{9}$$

Так как формулы (8) и (9) должны быть тождественны, то, сравнивая их, выводим:

$$a_{1,\,j+1}=\frac{1}{2+s},\quad b_{1,\,j+1}=\varphi(t_{j+1})+su_{1j}. \tag{10}$$

Пользуясь формулами (7) и (10), производя «прогонку» в прямом направлении (*прямой ход*), определяем две последовательности чисел:

$a_{1,\,j+1},\ a_{2,\,j+1},\ \ldots,\ a_{n-1,\,j+1}$ и $b_{1,\,j+1},\ b_{2,\,j+1},\ \ldots,\ b_{n-1,\,j+1}$.

Отсюда, применяя формулы (4) и (5), с помощью «*обратного хода*» находим значения искомой функции:

$$\left.\begin{aligned}
u_{n,\,j+1}&=\psi(t_{j+1}),\\
u_{n-1,\,j+1}&=(u_{n,\,j+1}+b_{n-1,\,j+1})\,a_{n-1,\,j+1},\\
u_{n-2,\,j+1}&=(u_{n-1,\,j+1}+b_{n-2,\,j+1})\,a_{n-2,\,j+1},\\
&\ldots\ldots\ldots\ldots\ldots\ldots\\
u_{1,\,j+1}&=(u_{2,\,j+1}+b_{1,\,j+1})\,a_{1,\,j+1}.
\end{aligned}\right\} \tag{11}$$

Таким образом, указан способ перехода от j-го слоя к $(j+1)$-му слою. Следовательно, отправляясь от известного начального (нулевого) слоя, можно шаг за шагом построить искомое решение $u(x,\ t)$ во всех точках сетки $(x_i,\ t_j)$.

Выясним устойчивость при $j\to+\infty$ конечно-разностной схемы (3). Пусть $v_{ij}=v(x_i,\ t_j)$ — решение уравнения (3), удовлетворяющее «возмущенным» условиям

$$v(x_i,\ 0)=f(x_i)+\varepsilon_{i0},\quad v(0,\ t_j)=\varphi(t_j),\quad v(l,\ t_j)=\psi(t_j),$$

где ε_{i0} — начальные ошибки. Погрешность

$$w_{ij}=v(x_i,\ t_j)-u(x_i,\ t_j)$$

удовлетворяет уравнению

$$w_{i-1,\,j+1}-(2+s)\,w_{i,\,j+1}+w_{i+1,\,j+1}+sw_{ij}=0 \tag{12}$$

и граничным условиям

$$w_{i0}=\varepsilon_{i0},\quad w_{0j}=w_{nj}=0. \tag{13}$$

Положим

$$w_{ij}^{(m)}=\lambda^{t_j}\sin\frac{m\pi x_i}{l}. \tag{14}$$

При целом m $(m=1,\ 2,\ 3,\ \ldots,\ n-1)$ функция $w_{ij}^{(m)}$ удовлетворяет вторым условиям (13).

Подставляя выражение (14) в уравнение (12), получим

$$\lambda^{t_j+k}\sin\frac{m\pi(x_i-h)}{l}-(2+s)\lambda^{t_j+k}\sin\frac{m\pi x_i}{l}+$$
$$+\lambda^{t_j+k}\sin\frac{m\pi(x_i+h)}{l}+s\lambda^{t_j}\sin\frac{m\pi x_i}{l}=0$$

или

$$\lambda^k\left[\sin\frac{m\pi(x_i-h)}{l}-2\sin\frac{m\pi x_i}{l}+\sin\frac{m\pi(x_i+h)}{l}\right]-$$
$$-s(\lambda^k-1)\sin\frac{\pi m x_i}{l}=0. \quad (15)$$

Так как

$$\sin\frac{m\pi(x_i-h)}{l}-2\sin\frac{m\pi x_i}{l}+\sin\frac{m\pi(x_i+h)}{l}=-4\sin\frac{m\pi x_i}{l}\sin^2\frac{m\pi h}{2l},$$

то, производя сокращение в формуле (15) на $\sin\frac{m\pi x_i}{l}$, будем иметь

$$4\lambda^k\sin^2\frac{m\pi h}{2l}+s(\lambda^k-1)=0.$$

Отсюда

$$\lambda^k=\frac{s}{s+4\sin^2\frac{m\pi h}{2l}}\leqslant 1, \quad (16)$$

если только $s>0$, и, следовательно,

$$|\lambda|\leqslant 1. \quad (17)$$

Таким образом, все решения $w_{ij}^{(m)}$ асимптотически стремятся к нулю при $j\to\infty$ или в крайнем случае, когда $\sin\frac{m\pi h}{2l}=0$, ограничены и при фиксированном x_i не возрастают по модулю. Так как решение w_{ij} представляет собой линейную комбинацию функций $w_{ij}^{(m)}$, то из неравенства (17) следует, что схема (3) устойчива при любом $s>0$.

Что касается ошибки аппроксимации для схемы (3), то эта ошибка, вообще говоря, есть $O(h^2+k)$ (ср. § 10).

Пример 1. Рассмотрим разобранную выше (§ 10) смешанную задачу для уравнения теплопроводности

$$\frac{\partial u}{\partial t}=\frac{\partial^2 u}{\partial x^2}$$

при начальных и краевых условиях

$$u(x,0)=4x(1-x),\quad u(0,t)=u(1,t)=0.$$

Примем $h=0{,}1$; $k=0{,}01$; следовательно, $s=\frac{h^2}{k}=1$. Полагая

$$u_{i,\,j+1}=a_{i,\,j+1}(b_{i,\,j+1}+u_{i+1,\,j+1}),$$

в силу формул (10) и (7) будем иметь:

$$a_{1,\,j+1}=1/3,\quad b_{1,\,j+1}=u_{1j}$$

и

$$a_{i,\,j+1}=\frac{1}{3-a_{i-1,\,j+1}},\quad b_{i,\,j+1}=a_{i-1,\,j+1}b_{i-1,\,j+1}+u_{ij}\quad (i>1).$$

В частности, при $j=0$ получаем формулы для первого слоя:

$$a_{1,1}=\frac{1}{3},\quad b_{1,1}=u_{10}, \tag{18}$$

$$a_{i,\,1}=\frac{1}{3-a_{i-1,1}},\quad b_{i,\,1}=a_{i-1,1}b_{i-1,1}+u_{i0}\quad (i>1). \tag{19}$$

Кроме того, на основании формул (11) имеем:

$$\left.\begin{aligned} &u_{n1}=0,\\ &u_{n-1,1}=(u_{n1}+b_{n-1,1})\,a_{n-1,1},\\ &u_{n-2,1}=(u_{n-1,1}+b_{n-2,1})\,a_{n-2,1},\\ &\ldots\ldots\ldots\ldots\ldots\ldots\\ &u_{11}=(u_{2,1}+b_{1,1})\,a_{1,1}. \end{aligned}\right\} \tag{20}$$

По формулам (18), (19) и (20) можно вычислить значения u_{ij} искомой функции для первого слоя ($j=1$). Полученные результаты приведены в таблице 72.

Таблица 72

Решение уравнения теплопроводности методом прогонки

i	0	1	2	3	4	5	6	7	8	9	10
u_{i0}	0	0,36	0,64	0,84	0,96	1,00	0,96	0,84	0,64	0,36	0
a_{i1}		0,333	0,375	0,381	0,382	0,382	0,382	0,382	0,382	0,382	
b_{i1}		0,360	0,760	1,125	1,389	1,530	1,544	1,430	1,186	0,813	
u_{i1}	0	0,310	0,572	0,764	0,882	0,921	0,882	0,764	0,571	0,310	0
u^*_{i1}	0	0,302	0,564	0,761	0,881	0,921	0,881	0,761	0,564	0,302	0

В последней строке таблицы 72 для сравнения приведены значения u^*_{i1} искомой функции, полученные обычным методом сеток при $h=\Delta x=1/10$ и $k=\Delta t=1/600$ (см. таблицу 71). Обращает на себя внимание значительное расхождение значений u_{i1} и u^*_{i1} вблизи границы области ($i=1$ и $i=9$). Это объясняется тем, что для таких точек примененные формулы численного дифференцирования обладают пониженной точностью. Для устранения этого неблагоприятного обстоятельства рекомендуется для точек, близких к границе области, использовать более точные формулы численного дифференцирования.

§ 13. Метод сеток для уравнений гиперболического типа

Остановимся на простейшем уравнении гиперболического типа, а именно уравнении свободных колебаний однородной ограниченной струны (см. § 1):

$$\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2}, \tag{1}$$

и будем искать решение уравнения (1) при заданных начальных и краевых условиях

$$u(x, 0) = f(x), \quad u_t(x, 0) = F(x) \quad (0 \leqslant x \leqslant l) \tag{2}$$

и

$$u(0, t) = \varphi(t), \quad u(l, t) = \psi(t) \quad (0 \leqslant t < \infty). \tag{3}$$

Решим эту смешанную задачу методом сеток [7], [8]. Как и в случае параболического уравнения, покроем полуполосу $0 \leqslant x \leqslant l$, $0 \leqslant t < \infty$ прямоугольной сеткой $x_i = ih, t_j = jk$ ($i = 0, 1, 2, \ldots, n$; $j = 0, 1, 2, \ldots$), где $\Delta x_i = x_{i+1} - x_i = h = l/n$ (n — целое) и $\Delta t_j = t_{j+1} - t_j = k$. На сетке x_i, t_j приближенно заменим дифференциальное уравнение (1) соответствующим конечно-разностным уравнением.

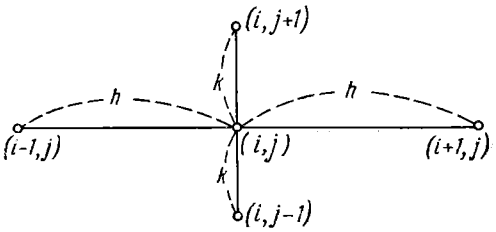

Рис. 75.

Пользуясь симметричными формулами для производных, будем иметь

$$\frac{u_{i,j+1} - 2u_{ij} + u_{i,j-1}}{k^2} = a^2 \frac{u_{i+1,j} - 2u_{ij} + u_{i-1,j}}{h^2}. \tag{4}$$

При $k = h/a$ уравнение (4) упрощается и принимает вид

$$u_{i,j+1} + u_{i,j-1} = u_{i+1,j} + u_{i-1,j},$$

откуда

$$u_{i,j+1} = u_{i+1,j} + u_{i-1,j} - u_{i,j-1}. \tag{5}$$

Из уравнения (5) видно, что для получения значений $u(x, t)$ в $(j+1)$-м слое используются значения $u(x, t)$ в двух предыдущих слоях: j-м и $(j-1)$-м (рис. 75). Для начала вычисления по формуле (5) также необходимо знать значения $u(x, t)$ на двух слоях, в то время как начальные условия (2) задают нам значения $u(x, t)$ лишь на нулевом слое $j = 0$. Однако, используя начальные условия, можно определить значения $u(x, t)$ на *фиктивном слое* с номером

$j=-1$. Для этого заменим производную во втором начальном условии конечно-разностным отношением. Тогда будем иметь

$$\frac{u_{i,-1}-u_{i0}}{-k}=F_i,$$

где $F_i=F(x_i)$. Отсюда

$$u_{i,-1}=u_{i0}-kF_i. \tag{6}$$

Теперь, зная значения $u(x, t)$ на слое $j=-1$, определяемые с помощью формулы (6), можно начать вычисления. Краевые условия (3) используются для получения значений u_{0j} и u_{nj}.

Вместо определения значений $u(x, t)$ на слое $j=-1$ можно вычислить значения $u(x, t)$ на слое $j=1$. Это достигается, например, с помощью формулы Тейлора

$$u_{i1}\approx u_{i0}+k\frac{\partial u_{i0}}{\partial t}+\frac{k^2}{2}\frac{\partial^2 u_{i0}}{\partial t^2}. \tag{7}$$

Учитывая, что согласно уравнению (1) имеем

$$\frac{\partial^2 u_{i0}}{\partial t^2}=a^2\frac{\partial^2 u_{i0}}{\partial x^2},$$

перепишем формулу (7) в другом виде, а именно:

$$u_{i1}\approx u_{i0}+k\frac{\partial u_{i0}}{\partial t}+\frac{a^2k^2}{2}\frac{\partial^2 u_{i0}}{\partial x^2}. \tag{8}$$

Из начальных условий (2), предполагая, что $f(x)\in C^{(2)}[0, l]$, получаем:

$$u_{i0}=f_i,\quad \frac{\partial u_{i0}}{\partial t}=F_i,\quad \frac{\partial^2 u_{i0}}{\partial x^2}=f_i''. \tag{9}$$

Подставляя эти значения в формулу (8), окончательно находим

$$u_{i1}\approx f_i+kF_i+\frac{a^2k^2}{2}f_i''. \tag{10}$$

Очевидно, формулу (10) целесообразно применять в том случае, когда функция $f(x)$ задана аналитическим выражением.

Пример 1. Методом сеток найти приближенное решение уравнения [6]

$$\frac{\partial^2 u}{\partial t^2}=\frac{\partial^2 u}{\partial x^2},$$

удовлетворяющее граничным и начальным условиям

$$u(0, t)=u(\pi, t)=0 \quad (0\leqslant t<\infty);$$
$$u(x, 0)=x(\pi-x), \qquad u_t(x, 0)=0 \quad (0\leqslant x\leqslant\pi).$$

Решение. В нашем случае $a=1$, поэтому $k=h$. Выбираем $k=h=\pi/18$. С помощью формулы (10) определим значения u_{i1}.

Так как

$$F_i=0, \quad f_i''=-2,$$

то

$$u_{i1}=u_{i0}-h^2=u_{i0}-0{,}03048.$$

Дальше решение проводится по формуле (5). Полученные значения приведены в таблице 73.

В таблице приведены лишь данные для $0\leqslant x\leqslant \pi/2$, так как график решения $u=u(x, t)$ симметричен относительно плоскости $x=\pi/2$.

Таблица 73

Решение уравнения колебаний струны методом сеток

t \ x	$x=0$	$x=h$	$x=2h$	$x=3h$	$x=4h$	$x=5h$	$x=6h$	$x=7h$	$x=8h$	$x=9h$
$t=0$	0	0,518	0,975	1,371	1,706	1,980	2,193	2,346	2,437	2,467
$t=h$	0	0,487	0,944	1,340	1,675	1,950	2,163	2,315	2,406	2,437
$t=2h$	0	0,426	0,853	1,249	1,584	1,858	2,071	2,224	2,315	2,346
$t=3h$	0	0,366	0,731	1,097	1,432	1,706	1,919	2,071	2,163	2,193
$t=4h$	0	0,305	0,609	0,914	1,218	1,493	1,706	1,858	1,950	1,980
$t=5h$	0	0,244	0,487	0,731	0,975	1,218	1,432	1,584	1,675	1,706

Замечание 1. Отметим одну особенность уравнения колебаний струны.

При решении задачи Коши для уравнения колебаний струны дифференциальный оператор

$$L[u]=\frac{\partial^2 u}{\partial t^2}-a^2\frac{\partial^2 u}{\partial x^2}$$

заменяется на сетке при условии, что $h=ak$, конечно-разностным оператором

$$L_h[u]=\frac{1}{h^2}(u_{i,j+1}-u_{i+1,j}-u_{i-1,j}+u_{i,j-1}). \tag{11}$$

Покажем, что в этом случае функция, являющаяся решением уравнения колебания струны, т. е. удовлетворяющая уравнению

$$L[u]=0, \tag{12}$$

является также решением уравнения

$$L_h[u]=0.$$

В самом деле, как известно [1]—[5], любое решение дифференциального уравнения (12) может быть представлено в виде

$$u(x, t)=\varphi(x-at)+\psi(x+at),$$

где φ и ψ — дважды дифференцируемые функции. Полагая

$$x_i = ih, \quad t_j = jk = \frac{jh}{a},$$

будем иметь

$$u_{ij} = u(x_i, t_j) = \varphi[(i-j)h] + \psi[(i+j)h].$$

Подставляя это выражение в формулу (11), получим

$$L_h[u] = \frac{1}{h^2}\{\varphi[(i-j-1)h] + \psi[(i+j+1)h] - \varphi[(i-j+1)h] - \\ - \psi[(i+j+1)h] - \varphi[(i-j-1)h] - \psi[(i+j-1)h] + \\ + \varphi[(i-j+1)h] + \psi[(i+j-1)h]\} \equiv 0.$$

Замечание 2. Если для уравнения колебаний струны (1) краевые условия (3) отсутствуют, то с помощью формулы (5) можно построить решение $u(x, t)$ соответствующей задачи Коши лишь в сеточной области плоскости Oxt, имеющей форму треугольника OAB (рис. 76), где OB и AB — характеристики

$$t = \frac{x}{a}, \qquad t = \frac{l-x}{a},$$

проходящие соответственно через точки $O(0, 0)$ и $A(l, 0)$ (см. § 2).

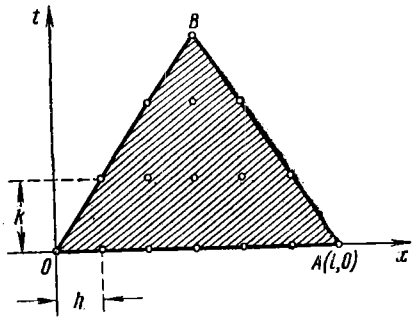

Рис. 76.

§ 14. Понятие о методе прямых

Метод прямых [21], [22] можно рассматривать как предельный случай метода сеток, когда при применении прямоугольной сетки один из линейных размеров ее стремится к нулю, а множество узлов в пределе заполняет некоторую систему прямолинейных параллельных отрезков. Идею этого метода мы изложим на примере линейного дифференциального уравнения эллиптического типа с двумя независимыми переменными и аналитическими коэффициентами.

Пусть в плоскости Oxy задана трапецевидная область G (рис. 77), основания которой лежат на прямых $y = \alpha$ и $y = \beta$ $(\alpha < \beta)$, а с боков эта область ограничена аналитическими кривыми

$$x = g_0(y) \ (L) \quad \text{и} \quad x = g_1(y) \ (\Gamma)$$
$$(\alpha \leqslant y \leqslant \beta; \ g_0(y) < g_1(y)),$$

причем область G целиком помещается в минимальном прямоугольнике $R\{a \leqslant x \leqslant b; \ \alpha \leqslant y \leqslant \beta\}$. В области G требуется, например, найти

решение $u = u(x, y)$ линейного эллиптического дифференциального уравнения

$$L[u] \equiv A(x, y)\frac{\partial^2 u}{\partial x^2} + 2B(x, y)\frac{\partial^2 u}{\partial x \partial y} + C(x, y)\frac{\partial^2 u}{\partial y^2} + \\ + a(x, y)\frac{\partial u}{\partial x} + b(x, y)\frac{\partial u}{\partial y} + c(x, y)u = f(x, y), \qquad (1)$$

удовлетворяющее на границе области G краевым условиям:

$$\left.\begin{aligned} &u(x, \alpha) = \varphi_0(x), \quad u(x, \beta) = \varphi_1(x), \\ &u(g_0(y), y) = \psi_0(y), \quad u(g_1(y), y) = \psi_1(y). \end{aligned}\right\} \qquad (2)$$

Будем предполагать, что коэффициенты и правая часть уравнения (1)

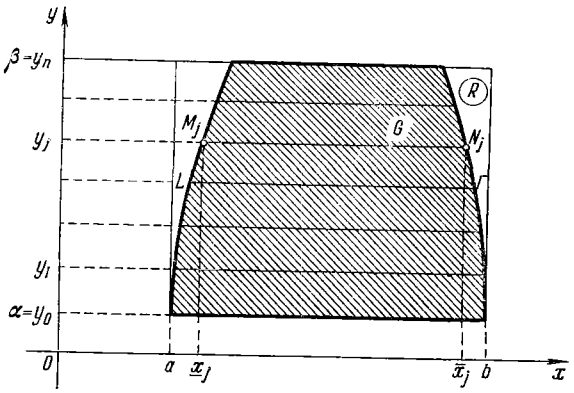

Рис. 77.

определены и аналитичны в замкнутом прямоугольнике R, причем выполнено *условие эллиптичности:*

$$D(x, y) = A(x, y)C(x, y) - B^2(x, y) > 0 \quad \text{при} \quad (x, y) \in R;$$

отсюда, в частности, следует, что

$$A(x, y) \neq 0, \; C(x, y) \neq 0 \; \text{при} \; (x, y) \in R. \qquad (3)$$

Допустим также, что функции $\varphi_0(x)$ и $\varphi_1(x)$ являются аналитическими на всем отрезке $[a, b]$, а функции $\psi_0(y)$ и $\psi_1(y)$ — аналитическими на отрезке $[\alpha, \beta]$ и, сверх того, выполнены условия согласованности:

$$\varphi_0(g_j(\alpha)) = \psi_j(\alpha), \quad \varphi_1(g_j(\beta)) = \psi_j(\beta), \qquad (j = 0, 1). \qquad (4)$$

Для получения по методу прямых приближенного решения краевой задачи (1) — (2) разделим отрезок $[\alpha, \beta]$ на n равных частей с помощью

точек $y_j = y_0 + jh$ $(y_0 = \alpha,\ y_n = \beta)$, $h = \frac{\beta - \alpha}{n}$ $(j = 0, 1, 2, \ldots, n)$ и через внутренние точки деления проведем семейство параллелей $y = y_j$ $(j = 1, 2, \ldots, n-1)$ На каждой такой прямой дифференциальное уравнение (1) приближенно заменим обыкновенным дифференциальным уравнением для искомых функций $u(x, y_j)$. Для этого в уравнении (1) избавимся от частного дифференцирования по y с помощью формул численного дифференцирования:

$$\left.\begin{aligned} &\left.\frac{\partial^2 u}{\partial y^2}\right|_{y=y_j} = \frac{1}{h^2}[u(x, y_{j+1}) - 2u(x, y_j) + u(x, y_{j-1})], \\ &\left.\frac{\partial^2 u}{\partial x\,\partial y}\right|_{y=y_j} = \frac{1}{2h}[u'_x(x, y_{j+1}) - u'_x(x, y_{j-1})], \\ &\left.\frac{\partial u}{\partial y}\right|_{y=y_j} = \frac{1}{2h}[u(x, y_{j+1}) - u(x, y_{j-1})] \\ &\qquad (j = 1, 2, \ldots, n-1). \end{aligned}\right\} \tag{5}$$

Введем сокращенные обозначения:

$$u(x, y_j) = u_j(x), \quad \frac{\partial u(x, y_j)}{\partial x} = u'_j(x), \quad \frac{\partial^2 u(x, y_j)}{\partial x^2} = u''_j(x),$$

а также

$$A(x, y_j) = A_j(x) \text{ и т. д.}$$

Тогда, подставляя выражения (5) в уравнение (1), получим следующую систему обыкновенных дифференциальных уравнений:

$$\begin{aligned} &A_j(x)\,u''_j(x) + \frac{B_j(x)}{h}[u'_{j+1}(x) - u'_{j-1}(x)] + \\ &\quad + \frac{C_j(x)}{h^2}[u_{j+1}(x) - 2u_j(x) + u_{j-1}(x)] + \\ &\quad + a_j(x)\,u'_j(x) + \frac{b_j(x)}{2h}[u_{j+1}(x) - u_{j-1}(x)] + c_j(x)\,u_j(x) = \\ &\qquad\qquad = f_j(x) \quad (j = 1, 2, \ldots, n-1). \end{aligned} \tag{6}$$

Кроме того, в силу краевых условий (2) имеем

$$u_0(x) = \varphi_0(x), \quad u_n(x) = \varphi_1(x) \tag{7}$$

и, следовательно,

$$u'_0(x) = \varphi'_0(x), \quad u'_n(x) = \varphi'_1(x).$$

Таким образом, от линейного дифференциального уравнения (1) с частными производными мы перешли к системе (6) из $n-1$ обыкновенных дифференциальных уравнений с $n-1$ неизвестными функциями $u_1(x)$, $u_2(x)$, $\ldots$, $u_{n-1}(x)$, где $u_0(x)$ и $u_n(x)$ известны и определяются формулами (7). Система (6) называется *системой уравнений метода прямых*.

Так как коэффициенты и правые части линейной системы (6) аналитичны, а следовательно, непрерывны на отрезке $[a, b]$ и старшие коэффициенты $A_j(x) \neq 0$ при $a \leqslant x \leqslant b$, то в силу известной теоремы из теории дифференциальных уравнений общее решение

$$u_j(x) = \varphi_j(x;\ C_1,\ C_2,\ \ldots,\ C_{2n-2}) \quad (j = 1,\ 2,\ \ldots,\ n-1)$$

системы (6) определено на отрезке $[a, b]$ и содержит $2n-2$ произвольных постоянных $C_1, C_2, \ldots, C_{2n-2}$, входящих в функции φ_i линейно.

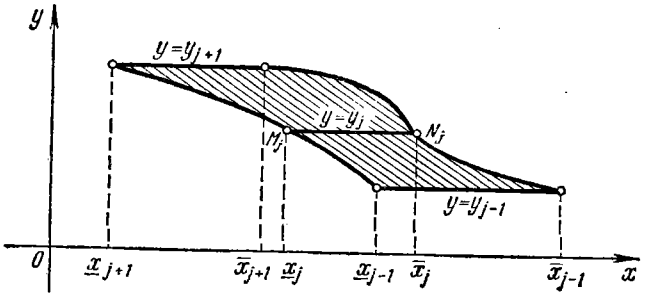

Рис. 78.

Для определения этих постоянных на основании краевых условий (2) получаем такое же число линейных алгебраических уравнений. А именно, пусть $\underline{x}_j = g_0(y_j)$ и $\bar{x}_j = g_1(y_j)$ — проекции на ось Ox концов отрезка M_jN_j, лежащего на параллели $y = y_j$ (см. рис. 77). Тогда на основании формул (2) имеем граничные условия

$$u_j(\underline{x}_j) = \psi_0(y_j), \quad u_j(\bar{x}_j) = \psi_1(y_j), \tag{8}$$

где $a \leqslant \underline{x}_j < \bar{x}_j \leqslant b \quad (j = 1,\ 2,\ \ldots,\ n-1)$.

Итак, наша задача (1)—(2) сводится к решению краевой задачи (6)—(8) для системы обыкновенных дифференциальных уравнений. Если краевая задача (6)—(8) имеет решение, то оно может быть найдено или обычным методом, когда система (6) допускает точное решение, или с помощью приближенных методов, описанных в гл. IV, причем функции $u_j(x)$ должны определяться на всем отрезке $[a, b]$.

После этого мы будем знать приближенные значения искомой функции $u(x, y)$ на семействе параллелей $y = y_j$ $(j = 0, 1, 2, \ldots, n)$. Значения функции $u(x, y)$ в промежуточных точках области G могут быть найдены интерполированием.

Отметим специфическую особенность краевой задачи (6)—(8): каждую из искомых функций $u_j(x)$ нужно определить на всем отрезке $[a, b]$, зная ее значения в двух, вообще говоря, внутренних точках $\underline{x}_j$ и $\bar{x}_j$ этого отрезка. Если мы найдем некоторую функцию $u_j(x)$ лишь при $\underline{x}_j \leqslant x \leqslant \bar{x}_j$, то это может оказаться недоста-

точным для решения задачи. Дело в том, что проекция на ось Ox отрезка M_jN_j параллели $y=y_j$ $(\underline{x}_j \leqslant x \leqslant \overline{x}_j)$ в общем случае не покрывает проекций на эту ось соседних параллелей $y=y_{j-1}$ $(\underline{x}_{j-1} \leqslant x \leqslant \overline{x}_{j-1})$ и $y=y_{j+1}$ $(\underline{x}_{j+1} \leqslant x \leqslant \overline{x}_{j+1})$ (рис. 78) и, значит, при этих неблагоприятных обстоятельствах для нахождения из системы (6) функций $u_{j-1}(x)$ или $u_{j+1}(x)$ нужно знать значения функции $u_j(x)$ вне отрезка $[\underline{x}_j, \overline{x}_j]$.

Мы рассмотрели простейший случай метода прямых, когда исходные данные аналитичны и, следовательно, допускают однозначное аналитическое продолжение. В случае лишь непрерывных коэффициентов уравнения (1) и непрерывных граничных условий (2) при применении метода прямых возникают дополнительные трудности, так как решение системы (6) уравнений метода прямых существенным образом зависит от значений ее коэффициентов и функций $u_0(x)$ и $u_n(x)$ в R вне области G, причем здесь это продолжение неоднозначно. Эти трудности отпадают, если область G представляет собой прямоугольник, стороны которого параллельны осям координат.

Если коэффициенты уравнения (1) зависят от x, то метод прямых приводит к системе линейных уравнений (6) с переменными коэффициентами, решение которой затруднительно. Поэтому здесь, вообще говоря, выгоднее использовать обычный метод сеток.

Если коэффициенты уравнения (1) не зависят от переменной x, то система (6) метода прямых состоит из линейных дифференциальных уравнений с постоянными коэффициентами. Способы решения таких систем хорошо разработаны, и метод прямых может оказаться тут сравнительно выгодным.

§ 15. Метод прямых для уравнения Пуассона

При замене линейного дифференциального уравнения второго порядка системой уравнений метода прямых (§ 14), вообще говоря, получается ошибка порядка h^2, где h — расстояние между прямыми. Эту ошибку можно уменьшить, если воспользоваться более точными формулами численного дифференцирования. Покажем это на примере уравнения Пуассона.

Пусть в прямоугольной области $R\{a \leqslant x \leqslant b;\ \alpha \leqslant y \leqslant \beta\}$ задано уравнение Пуассона

$$\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}=f(x, y) \tag{1}$$

и требуется найти решение $u=u(x, y)$ этого уравнения, удовлетворяющее краевым условиям:

$$\left.\begin{aligned} &u(x, \alpha)=\varphi_0(x), \quad u(x, \beta)=\varphi_1(x); \\ &u(a, y)=\psi_0(y), \quad u(b, y)=\psi_1(y), \end{aligned}\right\} \tag{2}$$

где функции f и ψ_k $(k=0,1)$ непрерывны и $\varphi_k \in C^{(2)}[a, b]$ $(k=0, 1)$.

Будем решать краевую задачу (1) — (2) методом прямых. Для этого выберем шаг $h=\frac{\beta-\alpha}{n}$ и через точки деления $y_j=y_0+hj$ ($j=0, 1, 2, \ldots, n$; $y_0=\alpha$, $y_n=\beta$) проведем параллели $y=y_j$. Пусть $u_j(x)=u(x, y_j)$. Предполагая, что функция $u(x, y)$ имеет непрерывные частные производные по y до шестого порядка включительно, разложим функции $u_{j+1}(x)=u(x, y_j+h)$ и $u_{j-1}(x)=u(x, y_j-h)$ по формуле Тейлора с точностью до $O(h^6)$. Имеем

$$u_{j+1}(x)=u_j(x)+\frac{\partial u(x, y_j)}{\partial y}h+\frac{\partial^2 u(x, y_j)}{\partial y^2}\cdot\frac{h^2}{2!}+\frac{\partial^3 u(x, y_j)}{\partial y^3}\cdot\frac{h^3}{3!}+$$
$$+\frac{\partial^4 u(x, y_j)}{\partial y^4}\cdot\frac{h^4}{4!}+\frac{\partial^5 u(x, y_j)}{\partial y^5}\cdot\frac{h^5}{5!}+\frac{\partial^6 u(x, y_j+\theta_1 h)}{\partial y^6}\cdot\frac{h^6}{6!}\quad(0<\theta_1<1)\qquad(3)$$

и

$$u_{j-1}(x)=u_j(x)-\frac{\partial u(x, y_j)}{\partial y}h+\frac{\partial^2 u(x, y_j)}{\partial y^2}\cdot\frac{h^2}{2}-$$
$$-\frac{\partial^3 u(x, y_j)}{\partial y^3}\cdot\frac{h^3}{3!}+\frac{\partial^4 u(x, y_j)}{\partial y^4}\cdot\frac{h^4}{4!}-$$
$$-\frac{\partial^5 u(x, y_j)}{\partial y^5}\cdot\frac{h^5}{5!}+\frac{\partial^6 u(x, y_j-\theta_2 h)}{\partial y^6}\cdot\frac{h^6}{6!}\quad(0<\theta_2<1).\qquad(4)$$

Сложив равенства (3) и (4), получим

$$u_{j+1}(x)-2u_j(x)+u_{j-1}(x)=\frac{\partial^2 u(x, y_j)}{\partial y^2}h^2+\frac{\partial^4 u(x, y_j)}{\partial y^4}\cdot\frac{h^4}{12}+O(h^6).\qquad(5)$$

Заменяя в формуле (5) функции

$$u_k(x)=u(x, y_k)\quad(k=j+1,\ j,\ j-1)$$

соответствующими вторыми производными $\frac{\partial^2 u(x, y_k)}{\partial y^2}$ и ограничиваясь членами порядка h^2, будем иметь

$$\frac{\partial^2 u(x, y_{j+1})}{\partial y^2}-2\frac{\partial^2 u(x, y_j)}{\partial y^2}+\frac{\partial^2 u(x, y_{j-1})}{\partial y^2}=\frac{\partial^4 u(x, y_j)}{\partial y^4}h^2+O(h^4).\qquad(6)$$

Исключая из формул (5) и (6) производную $\frac{\partial^4 u(x, y_j)}{\partial y^4}$ и отбрасывая члены порядка h^6, получим приближенную формулу

$$u_{j+1}(x)-2u_j(x)+u_{j-1}(x)=$$
$$=h^2\frac{\partial^2 u(x, y_j)}{\partial y^2}+\frac{h^2}{12}\left[\frac{\partial^2 u(x, y_{j+1})}{\partial y^2}-2\frac{\partial^2 u(x, y_j)}{\partial y^2}+\frac{\partial^2 u(x, y_{j-1})}{\partial y^2}\right],$$

которая после приведения подобных членов принимает вид

$$u_{j+1}(x)-2u_j(x)+u_{j-1}(x)=$$
$$=\frac{h^2}{12}\left[\frac{\partial^2 u(x, y_{j+1})}{\partial y^2}+10\frac{\partial^2 u(x, y_j)}{\partial y^2}+\frac{\partial^2 u(x, y_{j-1})}{\partial y^2}\right].\qquad(7)$$

Формула (7), имеющая точность $O(h^6)$, может быть использована для решения краевой задачи (1)—(2). Действительно, из уравнения (1) при $y=y_k$ имеем

$$\frac{\partial^2 u(x, y_k)}{\partial y^2} = f_k(x) - u_k''(x), \tag{8}$$

где $f_k(x) = f(x, y_k)$ $(k=1, 2, \ldots, n-1)$. Отсюда, заменяя в формуле (7) вторые частные производные по y их значениями из формулы (8), для определения решений $u_j(x)$ $(j=1, 2, \ldots, n-1)$ получим следующую систему обыкновенных дифференциальных уравнений:

$$u_{j+1}''(x) + 10u_j''(x) + u_{j-1}''(x) + \frac{12}{h^2}[u_{j+1}(x) - 2u_j(x) + u_{j-1}(x)] = \\ = f_{j+1}(x) + 10f_j(x) + f_{j-1}(x) \quad (j=1, 2, \ldots, n-1). \tag{9}$$

Эта усовершенствованная система (9) метода прямых была предложена М. Г. Слободянским [23] и аппроксимирует уравнение Пуассона с точностью до $\frac{12}{h^2} O(h^6) = O(h^4)$.

На основании краевых условий (2) дополнительно получаем:

$$u_0(x) = \varphi_0(x), \quad u_n(x) = \varphi_1(x); \quad u_j(a) = \psi_0(y_j), \quad u_j(b) = \psi_1(y_j) \quad (j=1, 2, \ldots, n-1). \tag{10}$$

Отсюда

$$u_0''(x) = \varphi_0''(x), \qquad u_n''(x) = \varphi_1''(x).$$

Общее решение системы (9), как известно, складывается из частного решения этой системы и общего решения соответствующей однородной системы

$$v_{j+1}''(x) + 10v_j''(x) + v_{j-1}''(x) + \\ + \frac{12}{h^2}[v_{j+1}(x) - 2v_j(x) + v_{j-1}(x)] = 0. \tag{11}$$

Очевидно, что общее решение системы (11) не зависит от области R и краевых условий (2) и для данного уравнения (1) может быть получено раз и навсегда. Приведем без доказательства (см. [22]) формулы общего решения системы (11):

$$v_j(x) = \sum_{k=1}^{n-1} \sin\frac{\pi k(y_j - y_0)}{l}(A_k e^{\sigma_j x} + B_k e^{-\sigma_j x}) \quad (j=1, 2, \ldots, n-1),$$

где $l = \beta - \alpha$, A_k и B_k — произвольные постоянные и

$$\sigma_j^2 = \frac{24}{h^2} \cdot \frac{\sin^2\frac{\pi(y_j - y_0)}{2l}}{5 + \cos\frac{\pi(y_j - y_0)}{l}}.$$

Частное решение неоднородной системы (9) находится обычным путем, в крайнем случае можно применить метод вариации произвольных постоянных. Для отыскания постоянных A_k и B_k на основании условий (10) получается алгебраическая система $2n-2$ уравнений.

Пример 1. В области $R\{0 \leqslant x \leqslant 3;\ 0 \leqslant y \leqslant 3\}$ задано уравнение Пуассона

$$\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}=x+y. \tag{12}$$

Методом прямых найти решение этого уравнения, удовлетворяющее однородным краевым условиям

$$u(0,\ y)=u(3,\ y)=u(x,\ 0)=u(x,\ 3)=0. \tag{13}$$

Решение. Примем $h=1$ и проведем прямые $y=1$ и $y=2$. Используя метод прямых, будем искать приближенное решение $u_j(x)=u(x,\ y_j)$ $(j=1,\ 2)$ задачи (12)—(13) на прямых $y=y_1$ и $y=y_2$, где $y_1=1$ и $y_2=2$. Выписывая систему (9), получим следующие два уравнения:

$$\left.\begin{aligned} u_2''(x)+10u_1''(x)+u_0''(x)+12\,[u_2(x)-2u_1(x)+u_0(x)]&= \\ =x+2+10\,(x+1)+(x+0),& \\ u_3''(x)+10u_2''(x)+u_1''(x)+12\,[u_3(x)-2u_2(x)+u_1(x)]&= \\ =x+3+10\,(x+2)+(x+1).& \end{aligned}\right\} \tag{14}$$

В силу краевых условий (13) имеем $u_0(x)=u_3(x)=0$, и, следовательно, $u_0''(x)=u_3''(x)=0$. Система (14) принимает следующий вид:

$$\left.\begin{aligned} 10u_1''(x)+u_2''(x)-24u_1(x)+12u_2(x)&=12x+12, \\ u_1''(x)+10u_2''(x)+12u_1(x)-24u_2(x)&=12x+24, \end{aligned}\right\} \tag{15}$$

причем

$$u_1(0)=u_1(3)=0,\quad u_2(0)=u_2(3)=0. \tag{16}$$

Соответствующая однородная система имеет вид

$$\left.\begin{aligned} 10v_1''(x)+v_2''(x)-24v_1(x)+12v_2(x)&=0, \\ v_1''(x)+10v_2''(x)+12v_1(x)-24v_2(x)&=0. \end{aligned}\right\} \tag{17}$$

Полагая

$$v_1(x)=Ae^{\lambda x},\quad v_2(x)=Be^{\lambda x} \tag{18}$$

и подставляя эти выражения в систему (17), после сокращения на $e^{\lambda x}$ получаем

$$\left.\begin{aligned} A\,(10\lambda^2-24)+B\,(\lambda^2+12)&=0, \\ A\,(\lambda^2+12)+B\,(10\lambda^2-24)&=0. \end{aligned}\right\} \tag{19}$$

Так как мы предполагаем, что решение (18) ненулевое, то определитель линейной системы (19) должен быть равен нулю. Отсюда получаем *характеристическое уравнение*

$$\begin{vmatrix} 10\lambda^2-24 & \lambda^2+12 \\ \lambda^2+12 & 10\lambda^2-24 \end{vmatrix}=0$$

или

$$(10\lambda^2-24)^2-(\lambda^2+12)^2=0,$$

т. е.

$$(9\lambda^2-36)(11\lambda^2-12)=0. \tag{20}$$

Следовательно, характеристические корни будут

$$\lambda_1=2,\quad \lambda_2=-2,\quad \lambda_3=\sqrt{\frac{12}{11}},\quad \lambda_4=-\sqrt{\frac{12}{11}}.$$

Соответствующие постоянные A и B определяются из системы (19). Имеем

$$\frac{A}{\lambda^2+12}=-\frac{B}{10\lambda^2-24}=C;$$

отсюда

$$\frac{A_k}{16}=-\frac{B_k}{16}\qquad (k=1,2);\qquad \frac{A_k}{\frac{144}{11}}=\frac{B_k}{\frac{144}{11}}\qquad (k=3,4).$$

Можно принять $A_k=-B_k=C_k$ $(k=1,2)$, $A_k=B_k=C_k$ $(k=3,4)$. Таким образом,

$$\left.\begin{aligned} v_1(x)&=C_1e^{2x}+C_2e^{-2x}+C_3e^{x\sqrt{12/11}}+C_4e^{-x\sqrt{12/11}},\\ v_2(x)&=-C_1e^{2x}-C_2e^{-2x}+C_3e^{x\sqrt{12/11}}+C_4e^{-x\sqrt{12/11}}.\end{aligned}\right\} \tag{21}$$

Частное решение системы (15) ищем в виде

$$\tilde{u}_1(x)=Ax+B,\quad \tilde{u}_2(x)=Cx+D.$$

Подставляя эти выражения в систему (15), для определения постоянных A, B, C, D получаем систему

$$\left.\begin{aligned} -2A+C&=1, & -2B+D&=1,\\ A-2C&=1, & B-2D&=2.\end{aligned}\right\}$$

Отсюда находим: $A=C=-1$, $B=-4/3$, $D=-5/3$, и, значит,

$$\tilde{u}_1(x)=-\left(x+\frac{4}{3}\right),\quad \tilde{u}_2(x)=-\left(x+\frac{5}{3}\right). \tag{22}$$

На основании формул (21) и (22) общее решение системы (15) имеет вид

$$\left.\begin{aligned} u_1(x)&=C_1e^{2x}+C_2e^{-2x}+C_3e^{x\sqrt{12/11}}+C_4e^{-x\sqrt{12/11}}-\left(x+\frac{4}{3}\right),\\ u_2(x)&=-C_1e^{2x}-C_2e^{-2x}+C_3e^{x\sqrt{12/11}}+C_4e^{-x\sqrt{12/11}}-\left(x+\frac{5}{3}\right).\end{aligned}\right\} \tag{23}$$

Для определения постоянных C_1, C_2, C_3, C_4 используем граничные условия (16). В силу этих условий и формул (23) получаем систему

$$\left.\begin{aligned} C_1+C_2+C_3+C_4&=\frac{4}{3},\\ -C_1-C_2+C_3+C_4&=\frac{5}{3},\\ C_1e^6+C_2e^{-6}+C_3e^{3\sqrt{12/11}}+C_4e^{-3\sqrt{12/11}}&=\frac{13}{3},\\ -C_1e^6-C_2e^{-6}+C_3e^{3\sqrt{12/11}}+C_4e^{-3\sqrt{12/11}}&=\frac{14}{3}.\end{aligned}\right\} \tag{24}$$

Отсюда

$$C_1=-\frac{1}{6}\cdot\frac{1}{e^6+1}=-\frac{1}{12}\cdot\frac{e^{-3}}{\operatorname{ch}3},\quad C_2=-\frac{1}{6}\cdot\frac{e^6}{e^6+1}=-\frac{1}{12}\cdot\frac{e^3}{\operatorname{ch}3},$$

$$C_3=\frac{3}{4}\cdot\frac{3-e^{-3\sqrt{12/11}}}{\operatorname{sh}(3\sqrt{12/11})},\quad C_4=\frac{3}{4}\cdot\frac{e^{3\sqrt{12/11}}-3}{\operatorname{sh}(3\sqrt{12/11})}.$$

Подставляя эти значения в формулы (23), после несложных упрощений окончательно находим

$$\left.\begin{aligned}
u(x,\ 1)&=-\frac{1}{6}\cdot\frac{\operatorname{ch}(2x-3)}{\operatorname{ch}3}+\\
&+\frac{3}{2}\cdot\frac{3\operatorname{sh}(x\sqrt{12/11})-\operatorname{sh}[(x-3)\sqrt{12/11}]}{\operatorname{sh}(3\sqrt{12/11})}-\left(x+\frac{4}{3}\right),\\
u(x,\ 2)&=\frac{1}{6}\cdot\frac{\operatorname{ch}(2x-3)}{\operatorname{ch}3}+\\
&+\frac{3}{2}\cdot\frac{3\operatorname{sh}(x\sqrt{12/11})-\operatorname{sh}[(x-3)\sqrt{12/11}]}{\operatorname{sh}(3\sqrt{12/11})}-\left(x+\frac{5}{3}\right).
\end{aligned}\right\}\quad(25)$$

Мы рассмотрели некоторые методы численного решения различных типов дифференциальных уравнений в частных производных.

Кроме рассмотренных разработаны также и другие методы, например метод Ричардсона [24], который обладает тем достоинством, что относительно мало загружает память машины.

ЛИТЕРАТУРА К ГЛАВЕ V

[1] Тихонов А. Н. и Самарский А. А., Уравнения математической физики, «Наука», 1964, гл. I, IV.

[2] Соболев С. Л., Уравнения математической физики, изд. 4, «Наука», 1966, лекции I—IV.

[3] Петровский И. Г., Лекции об уравнениях с частными производными, изд. 3, Физматгиз, 1961, гл. I и III.

[4] Кошляков Н. С., Основные дифференциальные уравнения математической физики, изд. 4, ОНТИ, 1936, гл. I.

[5] Смирнов В. И., Курс высшей математики, изд. 18, т. II, Физматгиз, 1962, гл. VII.

[6] Коллатц Л., Численные методы решения дифференциальных уравнений, ИЛ, 1953, гл. III и IV.

[7] Милн В. Э., Численное решение дифференциальных уравнений, ИЛ, 1955, гл. VIII.

[8] Рябенький В. С. и Филиппов А. Ф., Об устойчивости разностных уравнений, Гостехиздат, 1956, гл. I, II.

[9] Панов Д. Ю., Справочник по численному интегрированию дифференциальных уравнений в частных производных, «Наука», 1966.

[10] Саульев В. К., Интегрирование уравнений параболического типа методом сеток, Физматгиз, 1960.

[11] Современная математика для инженеров, под ред. Беккенбаха Э. Ф., гл. II, Браун Дж. В., Методы Монте-Карло, ИЛ, 1958.

[12] Демидович Б. П. и Марон И. А., Основы вычислительной математики, изд. 3, Гостехиздат, 1951, гл. XVII.
[13] Хаусхолдер А. С., Основы численного анализа, ИЛ, 1956, гл. VIII.
[14] Морз Ф. М. и Кимбелл Дж. Е., Методы исследования операций. Приложения, «Советское радио», 1956.
[15] Математика в СССР за сорок лет, т. I, Физматгиз, 1959, Гаврин М. К., Канторович Л. В., Приближенные и численные методы.
[16] Положий Г. Н. и др., Математический практикум, Физматгиз, 1960, гл. VII.
[17] Гутенмахер Л. И., Электрические модели, Изд. АН СССР, 1940.
[18] Кобринский Н. Е., Математические машины непрерывного действия, Гостехиздат, 1954.
[19] Китов А. И., Криницкий Н. А., Электронные цифровые машины и программирование, изд. 2, Физматгиз, 1961, гл. VIII.
[20] Гурса Э., Курс математического анализа, т. 3, ГТТИ, 1933, гл. XXVII.
[21] Михлин С. Г., Вариационные методы в математической физике, Гостехиздат, 1957, гл. XI.
[22] Березин И. С., Жидков Н. П., Методы вычислений, изд. 3, «Наука», 1966, т. II, гл. X.
[23] Слободянский М. Г., Способ приближенного интегрирования уравнений с частными производными и его применение к задачам теории упругости, Прикл. матем. и мех. 3, вып. I (1939).
[24] Ланс Дж. Н., Численные методы для быстродействующих вычислительных машин, ИЛ, 1962.

ГЛАВА VI

ВАРИАЦИОННЫЕ МЕТОДЫ РЕШЕНИЯ КРАЕВЫХ ЗАДАЧ

§ 1. Понятие о функционале и операторе

Введем сначала некоторые понятия функционального анализа, которые нам понадобятся в дальнейшем.

Определение 1. Пусть дан некоторый класс (множество) функций $K=\{y(x)\}$, где x — независимая переменная или совокупность нескольких независимых переменных $x=(x_1, x_2, \ldots, x_n)$. Говорят, что переменная величина $I=I[y(x)]$ есть *функционал* от функции $y(x)$ (функция от функции), если каждой функции $y(x)\in K$ по заданному закону ставится в соответствие определенное число I.

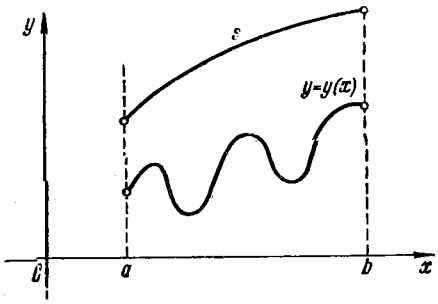

Рис. 79.

Класс функций $K=\{y(x)\}$, для которых определен данный функционал, называется *областью определения* или *областью задания* функционала, а сами функции называются *допустимыми*.

Пример 1. Пусть $K=\{y(x)\}$ — совокупность функций, дифференцируемых в точке $x=0$. Число $k=y'(0)$ можно рассматривать как функционал от $y(x)$, определенный в области K.

Пример 2. Рассмотрим множество K функций $y(x)$, непрерывно дифференцируемых на отрезке $[a, b]$, т. е. $y(x)\in C^{(1)}[a, b]$. Длина дуги s кривой $y=y(x)$ между точками $x=a$ и $x=b$ (рис. 79) есть функционал от $y(x)$ в области K, который может быть выражен формулой

$$s=\int_a^b \sqrt{1+y'^2}\,dx.$$

Пример 3. Пусть K — множество всех неотрицательных функций $z=f(x, y)$, непрерывных в замкнутой области $\overline{G}$ и обращающихся в нуль на ее границе Г (рис. 80). Объем

$$v=\iint\limits_G f(x, y)\,dx\,dy$$

есть функционал от $f(x, y)$.

Определение 2. Множество функций K называется *линейным* (или, короче, *линеалом*), если для каждых функций $u \in K$ и $v \in K$ сумма их $u+v \in K$, а также $\alpha u \in K$ (α — любая постоянная).

Например: а) множество полиномов, б) множество всех непрерывных функций, в) множество функций, обращающихся в нуль на границе области, и т. п. суть линеалы.

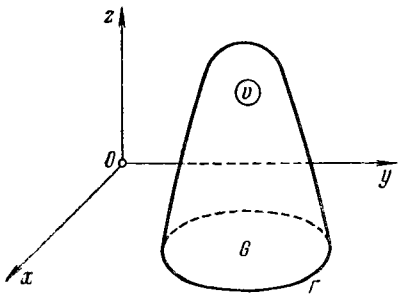

Рис. 80.

Определение 3. Функционал $I=I[y]$ называется *линейным*, если он определен на линейном множестве функций K и для любой пары допустимых функций u и v справедливо соотношение

$$I[\alpha u+\beta u]=\alpha I[u]+\beta I[v],$$

где α и β — произвольные постоянные.

Например, функционал $k=y'(0)$, рассмотренный в примере 1, является линейным.

Определение 4. Говорят, что на множестве $K=\{y(x)\}$ определен *оператор* L, если каждой функции $y(x) \in K$ по некоторому закону соответствует одна и только одна функция $z=z(x)$ *).

Это соответствие между функциями символически записывается следующим образом:

$$z=Ly \quad \text{или} \quad z=L(y).$$

Множество K функций $y=y(x)$, на котором определен данный оператор L, называется *областью задания* этого оператора, а функции $y \in K$ называются *допустимыми*.

Пример 4. Пусть $K=\{y(x)\}$ есть множество дифференцируемых функций. Тогда операцию $\frac{d}{dx}$ взятия производной можно рассматривать как

*) Возможно также, что функция $z(x)$ зависит от другой переменной $t=(t_1, \ldots, t_m)$.

оператор (так называемый *оператор дифференцирования*)

$$y'(x)=\frac{d}{dx}y(x).$$

В более общем случае, если $p_i(x)\in C[a, b]$ $(i=0, 1, 2, \ldots, n)$ и $y=y(x)\in C^{(n)}[a, b]$, то

$$z=Ly\equiv p_0(x)y^{(n)}+p_1(x)y^{(n-1)}+\ldots+p_n(x)y \tag{1}$$

есть оператор (*линейный дифференциальный оператор*), определенный на $K=C^{(n)}[a, b]$ со значениями $z\in C[a, b]$.

Действительно, для каждой допустимой функции y результат выполнения операций (1) есть некоторая непрерывная функция z. Например, если

$$L(y)=y''+y,$$

то $L(1)=1$; $L(x)=x$; $L(x^2)=2+x^2$; $L(e^x)=2e^x$; $L(\sin x)=0$ и т. д.

С помощью линейного дифференциального оператора Ly общее линейное дифференциальное уравнение с неизвестной функцией y можно коротко записать как

$$Ly=f(x), \tag{2}$$

где L — оператор вида (1) и $f(x)$ — известная непрерывная функция.

Пример 5. Рассмотрим множество функций $K=\{u(x, y)\}$ таких, что $u(x, y)\in C^{(2)}(G)$, где G — заданная область. Функция, определяемая формулой

$$\Delta u=\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2},$$

является оператором от u на множестве K (*оператор Лапласа*). Приравнивая этот оператор известной функции $f(x, y)$, получим *уравнение Пуассона*

$$\Delta u=f(x, y). \tag{3}$$

В частном случае, если $f(x, y)\equiv 0$, будем иметь *уравнение Лапласа*

$$\Delta u=0. \tag{4}$$

Таким образом, дифференциальные уравнения, обыкновенные и с частными производными, с более общей точки зрения можно рассматривать как операторные уравнения.

Определение 5. Оператор L называется *линейным*, если он определен на линейном множестве и для любой пары допустимых функций u и v линейная комбинация их $\alpha u+\beta v$ (α и β — произвольные постоянные) является также допустимой функцией, причем выполнены условия:

1) $L(\alpha u)=\alpha Lu$;

2) $L(u+v)=Lu+Lv$.

Отсюда вытекает, что

$$L(\alpha u+\beta v)=\alpha Lu+\beta Lv$$

при любых постоянных α и β.

Легко видеть, что операторы $\frac{d}{dx}y$, Ly и Δu, рассмотренные в примерах 4 и 5, являются линейными.

Пример 6. Оператор $Ly=y^2$ является нелинейным.

Действительно, $L(u+v)=(u+v)^2$ и $Lu+Lv=u^2+v^2$. Следовательно, $L(u+v)\neq Lu+Lv$, если только $uv\not\equiv 0$.

Пусть K есть множество функций $\{u\}$, определенных, действительных и непрерывных в области ω. Если $u\in K$ и $v\in K$, то число

$$(u,\ v)=\int\limits_{\omega} uv\,d\omega\ ^{*)}$$

называется *скалярным произведением функций* u и v. Очевидно, $(u,\ v)=(v,\ u)$.

Определение 6. Пусть линейный оператор L определен на линейном множестве функций u, заданных и непрерывных в области ω, и его значения Lu представляют собой функции, также определенные и непрерывные в ω. Тогда линейный оператор L называется *симметричным*, если для любых допустимых функций u и v справедливо соотношение

$$\int\limits_{\omega} vLu\,d\omega=\int\limits_{\omega} uLv\,d\omega,$$

т. е.

$$(Lu,\ v)=(u,\ Lv). \tag{5}$$

Если для любой допустимой функции u имеет место неравенство $(Lu,\ u)\geqslant 0$, причем $(Lu,\ u)=0$ тогда и только тогда, когда $u\equiv 0$, то оператор L называется *положительным*.

Пример 7. Рассмотрим оператор

$$Ly=-y'',$$

определенный на множестве функций $y\in C^{(2)}[0,\ 1]$ таких, что $y(0)=0$ и $y'(1)=0$.

Если u и v—допустимые функции, то получаем

$$\int\limits_0^1 (vLu-uLv)\,dx=\int\limits_0^1 (-vu''+uv'')\,dx=(uv'-vu')\Big|_0^1=0;$$

поэтому

$$\int\limits_0^1 vLu\,dx=\int\limits_0^1 uLv\,dx,$$

---

*) Если u и v—функции нескольких переменных, то интеграл $\int\limits_{\omega}$ является кратным.

т. е. $(Lu, v)=(u, Lv)$, и, таким образом, оператор L является симметричным.

Кроме того, учитывая, что $u\equiv 0$ есть в силу граничных условий единственная допустимая функция такая, что $u'\equiv 0$, при $u\not\equiv 0$ имеем

$$(Lu, u)=\int_0^1 uLu\,dx=-\int_0^1 uu''\,dx=-uu'\Big|_0^1+\int_0^1 u'^2\,dx>0,$$

причем $(Lu, u)=0$, если $u\equiv 0$. Следовательно, оператор L положителен.

§ 2. Вариационная задача

Пусть дан функционал

$$I=I[y(x)], \tag{1}$$

определенный на некотором множестве $K=\{y(x)\}$. Задача об отыскании экстремумов функционала (1) называется *вариационной задачей*. Более точно вариационная задача ставится следующим образом: требуется найти функцию $\bar{y}=\bar{y}(x)\in K$ такую, что для всех допустимых функций $y=y(x)$, достаточно близких к функции $\bar{y}(x)$, имеет место неравенство $I[y]\geqslant I[\bar{y}]$ в случае минимума или неравенство $I[y]\leqslant I[\bar{y}]$ в случае максимума. Заметим, что расстояние между функциями y и $\bar{y}$ можно понимать поразному.

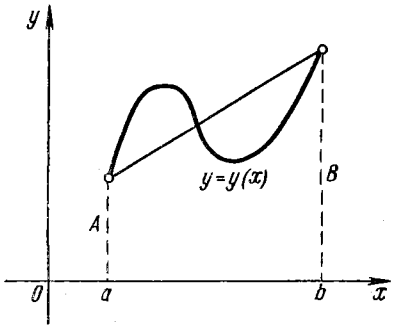

Рис. 81.

Пример 1. Рассмотрим задачу: среди гладких кривых $y=y(x)$, проходящих через точки $M(a, A)$ и $N(b, B)$, найти линию с наименьшей длиной дуги (рис. 81).

Задача сводится к нахождению минимума функционала

$$s=\int_a^b \sqrt{1+y'^2}\,dx$$

для кривых $y=y(x)$, принадлежащих классу $C^{(1)}[a, b]$ и таких, что $y(a)=A$ и $y(b)=B$. Из геометрических соображений очевидно, что искомым решением будет прямая

$$\bar{y}=A+\frac{B-A}{b-a}(x-a),$$

причем $s_{\min}=\sqrt{(b-a)^2+(B-A)^2}$.

§ 3. Основные теоремы вариационного метода решения краевых задач

Пусть в области G с границей Γ дано линейное дифференциальное уравнение с непрерывными коэффициентами (обыкновенное или с частными производными) и требуется найти решение y этого уравнения, удовлетворяющее на границе Γ заданным линейным однородным (краевым) условиям. Левую часть этого уравнения можно рассматривать как линейный оператор L, определенный на множестве K функций, обладающих непрерывными производными нужного порядка в $G+\Gamma$ и удовлетворяющих данным краевым условиям на Γ. Таким образом, наша краевая задача сводится к решению операторного уравнения

$$Ly=f(P), \tag{1}$$

где P обозначает совокупность независимых переменных, $f(P)$ — известная функция (которую мы будем считать непрерывной) и $y \in K$, причем функция y на границе Γ удовлетворяет краевым условиям

$$R[y]=0, \tag{2}$$

где R — известный линейный функционал или оператор более низкого порядка.

Заметим, что неоднородная краевая задача

$$Ly=f(P) \tag{3}$$

и

$$R[y]=\varphi(P) \quad \text{при} \quad P \in \Gamma, \tag{4}$$

где $\varphi(P)$ — известная функция, сводится к однородной, если положить $y=z+y_1$, где z — новая неизвестная функция и y_1 — достаточно гладкая функция, удовлетворяющая краевому условию (4): $R[y_1]=\varphi(P)$.

Действительно, из формул (3) и (4) получаем $Lz=f(P)-Ly_1$ и $R[z]=0$. Функцию y_1 обычно нетрудно найти подбором.

Идея вариационного метода применительно к нашему случаю состоит в том, что краевая задача (1) — (2) заменяется равносильной задачей об отыскании функции, дающей экстремум (обычно минимум) некоторому функционалу. Вариационный метод решения краевых задач получил широкое распространение после того, как немецкий математик Ритц в 1908 г. предложил удобный прием для построения приближенного решения вариационной задачи. Метод Ритца будет изложен в § 7.

Приведем две важные для дальнейшего теоремы.

Теорема 1. *Пусть L — симметричный линейный оператор, определенный и положительный в классе K. Тогда операторное уравнение* (1) *при наличии краевого условия* (2) *в классе K не может иметь*

двух решений, т. е. если существует решение краевой задачи (1)—(2), *то оно единственно.*

Доказательство. Предположим, что краевая задача (1)—(2) имеет два решения y_1 и y_2, т. е.

$$Ly_1 = f(P), \quad R[y_1] = 0 \tag{5}$$

и

$$Ly_2 = f(P), \quad R[y_2] = 0. \tag{6}$$

Вычитая из уравнений (5) соответствующие уравнения (6), в силу линейности оператора L и функционала R получим

$$L(y_1 - y_2) = 0, \quad R[y_1 - y_2] = 0, \tag{7}$$

т. е. $(y_1 - y_2) \in K$.

Умножая первое из полученных равенств скалярно на разность $y_1 - y_2$, будем иметь

$$(L(y_1 - y_2), (y_1 - y_2)) = 0. \tag{8}$$

Так как по условию оператор L положительный в классе K и функция $(y_1 - y_2) \in K$, то из формулы (8) следует $y_1 - y_2 \equiv 0$, т. е. $y_1 \equiv y_2$, что и требовалось доказать.

Теорема 2. *Пусть L — симметричный оператор, определенный и положительный в классе K, а $F[y]$ — функционал вида*

$$F[y] = (Ly, y) - 2(f, y) \equiv \int_\omega (Ly - 2f)\, y\, d\omega, \tag{9}$$

где $f = f(P)$ — правая часть уравнения (1).

Если краевая задача (1)—(2) *с однородными граничными условиями имеет решение $\bar{y}$, то это решение дает минимум функционалу $F[y]$.*

Обратно, если в классе K существует функция $\bar{y}$, дающая минимум функционалу (9), *то эта функция является решением уравнения* (1).

Доказательство. 1° Пусть $\bar{y}$ есть решение краевой задачи (1)—(2), т. е. $L\bar{y} = f(P)$ и $R[\bar{y}] = 0$. Заменяя $f(P)$ через $L\bar{y}$ в формуле (9), получим

$$F[y] = (Ly, y) - 2(L\bar{y}, y). \tag{10}$$

Пользуясь симметричностью оператора L, будем иметь

$$(L\bar{y}, y) = (\bar{y}, Ly) = (Ly, \bar{y}).$$

Поэтому

$$\begin{aligned} F[y] &= (Ly, y) - (Ly, \bar{y}) - (L\bar{y}, y) = \\ &= (Ly, y - \bar{y}) - [(L\bar{y}, y) - (L\bar{y}, \bar{y})] - (L\bar{y}, \bar{y}) = \\ &= (Ly, y - \bar{y}) - (L\bar{y}, y - \bar{y}) - (L\bar{y}, \bar{y}) = \\ &= (L(y - \bar{y}), (y - \bar{y})) - (L\bar{y}, \bar{y}). \end{aligned} \tag{11}$$

В правой части формулы (11) только первое слагаемое является переменным. Очевидно, $(y-\bar{y}) \in K$, поэтому в силу положительности оператора L имеем

$$(L(y-\bar{y}),\ (y-\bar{y})) \geqslant 0.$$

Следовательно, функционал $F[y]$ достигает своего наименьшего значения для тех и только тех допустимых функций y, для которых имеет место равенство

$$(L(y-\bar{y}),\ (y-\bar{y})) = 0.$$

Отсюда на основании определения положительного оператора получаем $y-\bar{y} \equiv 0$, т. е. $y \equiv \bar{y}$.

Заметим, что из формулы (11) следует, что наименьшее значение функционала $F[y]$ равно $F_{\min}(y) = F[\bar{y}] = -(L\bar{y}, \bar{y})$.

2° Пусть существует функция $\bar{y}$ из класса K, дающая минимум функционалу (9). Это значит, что для любой функции $y_1 \in K$ и достаточно близкой к $\bar{y}$ справедливо неравенство

$$F[y_1] - F[\bar{y}] \geqslant 0.$$

Положим $\eta = (y_1 - \bar{y}) \in K$ и рассмотрим семейство функций

$$y = \bar{y} + \alpha\eta, \tag{12}$$

где α — числовой параметр. Очевидно, при любом α функции y являются допустимыми и при достаточно малом $|\alpha|$ выполнено неравенство

$$\Delta F = F[y] - F[\bar{y}] \geqslant 0.$$

На основании формулы (9), выполняя тождественные преобразования, имеем

$$\begin{aligned}\Delta F &= (Ly, y) - 2(f, y) - (L\bar{y}, \bar{y}) + 2(f, \bar{y}) = \\ &= (Ly, y) - 2(L\bar{y}, y) + \\ &\quad + 2(L\bar{y} - f, y) - 2(L\bar{y} - f, \bar{y}) + (L\bar{y}, \bar{y}) = \\ &= (Ly, y) - 2(L\bar{y}, y) + 2(L\bar{y} - f, y - \bar{y}) + (L\bar{y}, \bar{y}) \geqslant 0.\end{aligned} \tag{13}$$

Отсюда, используя преобразование (11) и формулу (12), находим

$$\begin{aligned}\Delta F &= (L(y-\bar{y}),\ (y-\bar{y})) - (L\bar{y}, \bar{y}) + 2(L\bar{y} - f, y - \bar{y}) + (L\bar{y}, \bar{y}) = \\ &= \alpha^2 (L\eta, \eta) + 2\alpha (L\bar{y} - f, \eta) \geqslant 0.\end{aligned} \tag{14}$$

Левая часть неравенства (14) представляет собой квадратный трехчлен относительно параметра α, причем этот трехчлен не может менять знака. Следовательно, соответствующее квадратное уравнение заведомо не имеет действительных различных корней и, значит, обладает неположительным дискриминантом, т. е.

$$(L\bar{y} - f, \eta)^2 - (L\eta, \eta) \cdot 0 \leqslant 0;$$

отсюда $(L\bar{y}-f, \eta)=0$. Таким образом,

$$\int_{\omega} (L\bar{y}-f)\,\eta\, d\omega = 0 \tag{15}$$

для любой функции $\eta \in K$. В силу произвольности функции η отсюда следует (см. [3]), что $L\bar{y}-f\equiv 0$, т. е. $L\bar{y}=f$. Таким образом, $\bar{y}$ есть решение нашей краевой задачи.

§ 4. Сведение линейной краевой задачи для обыкновенного дифференциального уравнения второго порядка к вариационной задаче

Рассмотрим обыкновенное дифференциальное уравнение

$$y''+P(x)\,y'+Q(x)\,y=\Phi(x) \tag{1}$$

с линейными краевыми условиями

$$\left.\begin{aligned} \alpha_1 y'(a)+\alpha y(a)=A,\\ \beta_1 y'(b)+\beta y(b)=B,\end{aligned}\right\} \tag{2}$$

где функции $P(x)$, $Q(x)$ и $\Phi(x)$ непрерывны на отрезке $[a, b]$ и

$$|\alpha_1|+|\alpha|\neq 0, \quad |\beta_1|+|\beta|\neq 0.$$

Приведем уравнение (1) к специальному, так называемому *самосопряженному виду*. Для этого умножим все его члены на положительный множитель

$$p(x)=e^{\int_a^x P(x)\,dx},$$

после чего получим

$$p(x)\,y''(x)+p(x)\,P(x)\,y'+p(x)\,Q(x)\,y=p(x)\,\Phi(x). \tag{3}$$

Так как

$$p'(x)=e^{\int_a^x P(x)\,dx}\,P(x)=p(x)\,P(x),$$

то уравнение (3) можно записать в виде

$$\frac{d}{dx}[p(x)\,y']+q(x)\,y=f(x), \tag{4}$$

где $p(x)>0$, $q(x)=p(x)\,Q(x)$, $f(x)=p(x)\,\Phi(x)$.

Дифференциальное уравнение второго порядка вида (4) называется *самосопряженным*. Вводя линейный оператор

$$Ly=-\frac{d}{dx}[p(x)\,y']-q(x)\,y, \tag{5}$$

получим

$$Ly = -f(x), \tag{6}$$

где $p(x)$, $p'(x)$, $q(x)$ и $f(x)$ непрерывны на отрезке $[a, b]$.

Предположим сначала, что краевые условия (2) являются однородными, т. е.

$$\alpha_1 y'(a) + \alpha y(a) = 0, \quad \beta_1 y'(b) + \beta y(b) = 0, \tag{7}$$

где $|\alpha_1| + |\alpha| \neq 0$ и $|\beta_1| + |\beta| \neq 0$, причем без нарушения общности рассуждений можно предполагать, что $\alpha_1 \geqslant 0$ и $\beta_1 \geqslant 0$.

Покажем, что в этом случае оператор L является самосопряженным (симметричным) в классе функций $K = \{y\}$, непрерывных на отрезке $[a, b]$ вместе со своими первыми и вторыми производными ($y \in C^{(2)}[a, b]$) и удовлетворяющих на концах отрезка $[a, b]$ однородным краевым условиям (7).

Пусть $u \in K$ и $v \in K$. На основании формулы (5) имеем

$$(Lu, v) - (Lv, u) =$$

$$= \int_a^b \left\{ -v\left[\frac{d}{dx}(p(x)u') + q(x)u\right] + u\left[\frac{d}{dx}(p(x)v') + q(x)v\right]\right\} dx =$$

$$= \int_a^b [p(x)(uv'' - vu'') + p'(x)(uv' - vu')]\, dx =$$

$$= \int_a^b \frac{d}{dx}[p(x)(uv' - vu')]\, dx =$$

$$= p(x)(uv' - vu')\Big|_a^b = p(b)w(b) - p(a)w(a), \tag{8}$$

где

$$w(x) = \begin{vmatrix} u(x) & v(x) \\ u'(x) & v'(x) \end{vmatrix}. \tag{9}$$

Функции $u = u(x)$ и $v = v(x)$ удовлетворяют однородным краевым условиям

$$\alpha_1 u'(a) + \alpha u(a) = 0, \quad \alpha_1 v'(a) + \alpha v(a) = 0,$$

где $\alpha_1 \neq 0$ или $\alpha \neq 0$. Следовательно,

$$u'(a) = -\frac{\alpha}{\alpha_1} u(a), \quad v'(a) = -\frac{\alpha}{\alpha_1} v(a)$$

или

$$u(a) = -\frac{\alpha_1}{\alpha} u'(a), \quad v(a) = -\frac{\alpha_1}{\alpha} v'(a).$$

Поэтому $w(a) = 0$. Аналогично доказывается, что $w(b) = 0$.

Следовательно, из формулы (8) вытекает $(Lu, v)-(Lv, u)=0$, и, значит, $(Lu, v)=(Lv, u)$, т. е. оператор L симметричен.

Выясним, при каких условиях оператор L является положительным. Для функции $y \in K$ имеем

$$(Ly, y)=-\int_a^b \left\{\frac{d}{dx}[p(x)y']+q(x)y\right\} y\,dx. \tag{10}$$

Интегрируя по частям первый член формулы (10), получим

$$(Ly, y)=-p(x)\,yy'\Big|_a^b+\int_a^b [p(x)y'^2-q(x)y^2]\,dx. \tag{11}$$

Так как $p(x)>0$, то из формулы (11) вытекает, что оператор L положителен, если

$$q(x)\leqslant 0 \quad \text{при} \quad a\leqslant x\leqslant b, \tag{12}$$

$$y(a)y'(a)\geqslant 0, \quad y(b)y'(b)\leqslant 0. \tag{13}$$

Так как $\alpha_1\geqslant 0$ и $\beta_1\geqslant 0$, то в силу краевых условий (7) неравенства (13) эквивалентны неравенствам

$$\alpha\leqslant 0, \quad \beta\geqslant 0. \tag{14}$$

Таким образом, краевая задача (6)—(7) при наличии неравенств (12) и (14), согласно теореме 2 из § 3, равносильна задаче о минимуме функционала

$$F[y]=(Ly, y)+2(f, y) \tag{15}$$

в классе функций K. Используя формулу (11), имеем

$$F[y]=p(a)y(a)y'(a)-p(b)y(b)y'(b)+ \\ +\int_a^b [p(x)y'^2-q(x)y^2+2f(x)y]\,dx. \tag{16}$$

В частности, если $\alpha_1>0$ и $\beta_1>0$, то получим

$$F[y]=-\frac{\alpha}{\alpha_1}p(a)y^2(a)+\frac{\beta}{\beta_1}p(b)y^2(b)+ \\ +\int_a^b [p(x)y'^2-q(x)y^2+2f(x)y]\,dx. \tag{17}$$

Аналогичные выражения получаем для других случаев.

Рассмотрим теперь краевую задачу (6) с неоднородными краевыми условиями (2) в предположении, что выполнены неравенства (12) и (14). Оператор L в классе функций K_1, удовлетворяющих условиям (2), вообще говоря, не является симметричным и положитель-

ным, поэтому нельзя непосредственно использовать теорему 2 предыдущего параграфа.

Пусть $z = z(x) \in C^{(2)}[a, b]$ и удовлетворяет условиям (2), т. е.

$$\alpha_1 z'(a) + \alpha z(a) = A, \quad \beta_1 z'(b) + \beta z(b) = B. \tag{18}$$

Обозначая через y решение краевой задачи (6), (2), введем функцию $u = u(x)$, определяемую равенством

$$u = y - z. \tag{19}$$

Функция u удовлетворяет однородным краевым условиям

$$\alpha_1 u'(a) + \alpha u(a) = 0, \qquad \beta_1 u'(b) + \beta u(b) = 0 \tag{20}$$

и является решением уравнения $Lu = Ly - Lz$, т. е.

$$Lu = -f(x) - Lz. \tag{21}$$

Таким образом, $u \in K$. Оператор Lu в классе функций K является симметричным и положительным, и, следовательно, решение u краевой задачи (20)—(21), в силу теоремы 2 из § 3, дает минимум функционалу

$$F[u] = (Lu, u) + 2(f(x) + Lz, u).$$

Отсюда на основании формулы (15) имеем

$$F[u] = p(a)\,u(a)\,u'(a) - p(b)\,u(b)\,u'(b) + \\ + \int_a^b [p(x)\,u'^2 - q(x)\,u^2 + 2(f(x) + Lz)\,u]\,dx. \tag{22}$$

Из равенства (19) получаем, что решение y краевой задачи (6), (2) дает минимум функционалу

$$\begin{aligned} F_1[y] = {} & p(a)[y(a) - z(a)][y'(a) - z'(a)] - \\ & - p(b)[y(b) - z(b)][y'(b) - z'(b)] + \\ & + \int_a^b [p(x)(y' - z')^2 - q(x)(y - z)^2 + 2(f(x) + Lz)(y - z)]\,dx = \\ & = p(a)[y(a) - z(a)][y'(a) - z'(a)] - \\ & - p(b)[y(b) - z(b)][y'(b) - z'(b)] + \\ & + \int_a^b [p(x)\,y'^2 - q(x)\,y^2 + 2f(x)\,y]\,dx + \\ & + \int_a^b [p(x)\,z'^2 - q(x)\,z^2 - 2f(x)\,z]\,dz + \\ & + 2\int_a^b [-p(x)\,y'z' + q(x)\,yz + (y - z)\,Lz]\,dx. \end{aligned} \tag{23}$$

Используя интегрирование по частям, будем иметь

$$\int_a^b (y-z)\,Lz\,dx = -\int_a^b (y-z)\left[\frac{d}{dx}(p(x)z') + q(x)z\right]dx =$$

$$= -(y-z)\,p(x)\,z'\Big|_a^b + \int_a^b [p(x)z'(y'-z') - q(x)z(y-z)]\,dx =$$

$$= p(a)[y(a)-z(a)]z'(a) - p(b)[y(b)-z(b)]z'(b) +$$

$$+\int_a^b [p(x)z'(y'-z') - q(x)z(y-z)]\,dx.$$

Внося это выражение в формулу (23), после несложных упрощений получим

$$F_1(y) = p(a)[y(a)-z(a)][y'(a)+z'(a)] -$$

$$-p(b)[y(b)-z(b)][y'(b)+z'(b)] +$$

$$+\int_a^b [p(x)y'^2 - q(x)y^2 + 2f(x)y]\,dx -$$

$$-\int_a^b [p(x)z'^2 - q(x)z^2 + 2f(x)z]\,dx. \qquad (24)$$

Пусть $\alpha_1 > 0$ и $\beta_1 > 0$. Из краевых условий (2) имеем

$$y'(a) = \frac{A-\alpha y(a)}{\alpha_1}, \quad z'(a) = \frac{A-\alpha z(a)}{\alpha_1}$$

и

$$y'(b) = \frac{B-\beta y(b)}{\beta_1}, \quad z'(b) = \frac{B-\beta z(b)}{\beta_1}.$$

Тогда

$$F_1[y] = \frac{p(a)}{\alpha_1}[2Ay(a) - \alpha y^2(a)] - \frac{p(b)}{\beta_1}[2By(b) - \beta y^2(b)] +$$

$$+\int_a^b [p(x)y'^2 - q(x)y^2 + 2f(x)y]\,dx -$$

$$-\left\{\frac{p(a)}{\alpha_1}[2Az(a) - \alpha z^2(a)] - \frac{p(b)}{\beta_1}[2Bz(b) - \beta z^2(b)] +\right.$$

$$\left.+\int_a^b [p(x)z'^2 - q(x)z^2 + 2f(x)z]\,dx\right\}. \qquad (25)$$

Так как стоящие в фигурной скобке члены формулы (25) фиксированы и не меняются при изменении функции y, то вместо

функционала $F_1[y]$ можно рассмотреть функционал

$$\Phi[y]=\frac{p(a)}{\alpha_1}[2Ay(a)-\alpha y^2(a)]-\frac{p(b)}{\beta_1}[2By(b)-\beta y^2(b)]+$$
$$+\int_a^b[p(x)y'^2-q(x)y^2+2f(x)y]\,dx. \quad (26)$$

Таким образом, краевая задача (6), (2) с неоднородными краевыми условиями в предположении, что имеют место неравенства (12) и (14), эквивалентна вариационной задаче для функционала (26) в классе функций K_1, удовлетворяющих заданным краевым условиям.

Замечание. 1° Если $\alpha_1=0$ и $\beta_1\neq 0$, то $y(a)=z(a)=A/\alpha$. Из формулы (24) вытекает, что за $\Phi[y]$ можно принять функционал

$$\Phi[y]=-\frac{p(b)}{\beta_1}[2By(b)-\beta y^2(b)]+$$
$$+\int_a^b[p(x)y'^2-q(x)y^2+2f(x)y]\,dx.$$

2° Аналогично доказывается, что если $\alpha_1=0$ и $\beta_1=0$, то

$$\Phi[y]=\int_a^b[p(x)y'^2-q(x)y^2+2f(x)y]\,dx.$$

§ 5. Краевые задачи для уравнений Пуассона и Лапласа

Пусть дано уравнение Пуассона

$$-\Delta u=f(x,\ y), \quad (1)$$

где

$$\Delta u=\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2} \quad \text{и} \quad f(x,\ y)\in C(G).$$

Требуется найти решение уравнения (1), непрерывное в замкнутой области $\bar{G}=G+\Gamma$ и удовлетворяющее на границе Γ этой области краевому условию

$$u|_\Gamma=\varphi(P), \quad (2)$$

где $P=(x,\ y)$ и $\varphi(P)$ — заданная непрерывная функция. Предположим вначале, что $\varphi(P)\equiv 0$, т. е.

$$u|_\Gamma=0, \quad (3)$$

и будем решать однородную краевую задачу (1), (3). Покажем, что в классе функций $K=\{u(x)\}$, непрерывных в $\bar{G}$ вместе со своими первыми и вторыми производными и обращающихся на контуре Γ в нуль, оператор $Lu=-\Delta u$ симметричен и положителен.

Пусть $u \in K$ и $v \in K$. Составим выражение

$$(Lu, v) - (Lv, u) = \\ = \iint\limits_G \left[-v \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right) + u \left(\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} \right) \right] dx\, dy = \\ = \iint\limits_G \left[\frac{\partial}{\partial x} \left(u \frac{\partial v}{\partial x} - v \frac{\partial u}{\partial x} \right) + \frac{\partial}{\partial y} \left(u \frac{\partial v}{\partial y} - v \frac{\partial u}{\partial y} \right) \right] dx\, dy.$$

Применяя известную формулу Грина

$$\int\limits_\Gamma P\, dx + Q\, dy = \iint\limits_G \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx\, dy$$

и используя нулевые граничные условия $u|_\Gamma = 0$, $v|_\Gamma = 0$, получим

$$(Lu, v) - (Lv, u) = \\ = \int\limits_\Gamma \left[-\left(u \frac{\partial v}{\partial y} - v \frac{\partial u}{\partial y} \right) dx + \left(u \frac{\partial v}{\partial x} - v \frac{\partial u}{\partial x} \right) dy \right] = 0; \quad (4)$$

следовательно, $(Lu, v) = (Lv, u)$, и, значит, оператор L симметричен.

Далее установим положительность оператора L. Имеем

$$(Lu, u) = -\iint\limits_G u \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right) dx\, dy = \\ = -\iint\limits_G \left[\frac{\partial}{\partial x} \left(u \frac{\partial u}{\partial x} \right) + \frac{\partial}{\partial y} \left(u \frac{\partial u}{\partial y} \right) \right] dx\, dy + \iint\limits_G \left[\left(\frac{\partial u}{\partial x} \right)^2 + \left(\frac{\partial u}{\partial y} \right)^2 \right] dx\, dy.$$

Применяя к первому интегралу формулу Грина, в силу краевых условий для функции u получим

$$(Lu, u) = -\int\limits_\Gamma \left(-u \frac{\partial u}{\partial y} dx + u \frac{\partial u}{\partial x} dy \right) + \\ + \iint\limits_G \left[\left(\frac{\partial u}{\partial x} \right)^2 + \left(\frac{\partial u}{\partial y} \right)^2 \right] dx\, dy = \\ = \iint\limits_G \left[\left(\frac{\partial u}{\partial x} \right)^2 + \left(\frac{\partial u}{\partial y} \right)^2 \right] dx\, dy, \quad (5)$$

т. е.

$$(Lu, u) = (-\Delta u, u) \geqslant 0.$$

Если $(Lu, u) = 0$, то из формулы (5) следует, что

$$\frac{\partial u}{\partial x} = \frac{\partial u}{\partial y} = 0.$$

Отсюда $u(x, y) = c$, и на основании краевого условия (3) имеем

$$u(x, y) \equiv 0.$$

Следовательно, оператор L положителен.

Таким образом, для краевой задачи (1) с однородными краевыми условиями (3) выполнены условия теоремы 2 из § 3. Следовательно, эта задача эквивалентна вариационной задаче для функционала

$$F[u]=(Lu,\ u)-2(u,\ f) \tag{6}$$

в классе функций u, принадлежащих множеству K. В силу формулы (5) получаем

$$F[u]=\iint\limits_G \left[\left(\frac{\partial u}{\partial x}\right)^2+\left(\frac{\partial u}{\partial y}\right)^2-2fu\right] dx\,dy. \tag{7}$$

Рассмотрим теперь краевую задачу (1) с неоднородными краевыми условиями (2), и пусть $K_1=\{u(x,\ y)\}$ — класс функций $u\in C^{(2)}(G+\Gamma)$, удовлетворяющих условиям (2).

Следуя идее предыдущего параграфа, построим функцию $z=z(x,\ y)\in C^{(2)}(G+\Gamma)$, для которой выполнены краевые условия (2). Введем функцию

$$v(x,\ y)=u(x,\ y)-z(x,\ y), \tag{8}$$

где $u(x,\ y)$ — решение нашей неоднородной краевой задачи. Тогда функция $v=v(x,\ y)$ удовлетворяет на контуре Γ однородному краевому условию

$$v|_\Gamma=0 \tag{9}$$

и является решением уравнения

$$Lv=Lu-Lz=f(P)-Lz, \tag{10}$$

где $Lz=-\left(\frac{\partial^2 z}{\partial x^2}+\frac{\partial^2 z}{\partial y^2}\right)$ — известная функция. Функция $v=v(x,\ y)$, являясь решением однородной краевой задачи (10)—(9), на основании формулы (6) дает наименьшее значение функционалу

$$F[v]=(Lv,\ v)-2(v,\ f(P)-Lz). \tag{11}$$

Возвращаясь в последнем равенстве к функции u (см. (8)) и используя свойства скалярного произведения и линейного оператора L, получим

$$\begin{aligned}F[u-z]\equiv F_1[u]&=(L(u-z),\ u-z)-2(u-z,\ f(P)-Lz)=\\&=(Lu,\ u)-2(u,\ f)+(u,\ Lz)-(z,\ Lu)+2(z,\ f)-(Lz,\ z).\end{aligned} \tag{12}$$

Так как последние два члена формулы (12) не зависят от искомой функции $u=u(x,\ y)$, то функция $\bar u=\bar u(x,\ y)$, дающая наименьшее значение функционалу (12), будет минимизировать функционал

$$F_2[u]=(Lu,\ u)-2(u,\ f)+[(Lz,\ u)-(Lu,\ z)]. \tag{13}$$

Покажем, что функционал (13) можно заменить функционалом, не содержащим функцию z. Используя преобразование, примененное

в формуле (4), имеем

$$(Lz,\ u)-(Lu,\ z)=\iint\limits_G (z\,\Delta u-u\,\Delta z)\,dz=$$

$$=\int\limits_\Gamma\left[-\left(z\frac{\partial u}{\partial y}-u\frac{\partial z}{\partial y}\right)dx+\left(z\frac{\partial u}{\partial x}-u\frac{\partial z}{\partial x}\right)dy\right]=\int\limits_\Gamma\left(z\frac{\partial u}{\partial n}-u\frac{\partial z}{\partial n}\right)ds,$$

где n — внешняя нормаль к Γ и

$$\frac{\partial u}{\partial n}=\frac{\partial u}{\partial x}\frac{dy}{ds}-\frac{\partial u}{\partial y}\frac{dx}{ds},\qquad \frac{\partial z}{\partial n}=\frac{\partial z}{\partial x}\frac{dy}{ds}-\frac{\partial z}{\partial y}\frac{dx}{ds}.$$

Отсюда, так как $z|_\Gamma=u|_\Gamma=\varphi(x,\ y)$, получаем

$$(Lz,\ u)-(Lu,\ z)=\int\limits_\Gamma\varphi(x,\ y)\left(\frac{\partial u}{\partial n}-\frac{\partial z}{\partial n}\right)ds. \tag{14}$$

С другой стороны, на основании формулы (5) находим

$$(Lu,\ u)=-\int\limits_\Gamma u\frac{\partial u}{\partial n}ds+\iint\limits_G\left[\left(\frac{\partial u}{\partial x}\right)^2+\left(\frac{\partial u}{\partial y}\right)^2\right]dx\,dy=$$

$$=-\int\limits_\Gamma\varphi(x,\ y)\frac{\partial u}{\partial n}ds+\iint\limits_G\left[\left(\frac{\partial u}{\partial x}\right)^2+\left(\frac{\partial u}{\partial y}\right)^2\right]dx\,dy. \tag{15}$$

Подставляя (14) и (15) в формулу (13), будем иметь

$$F_2[u]=\iint\limits_G\left[\left(\frac{\partial u}{\partial x}\right)^2+\left(\frac{\partial u}{\partial y}\right)^2-2fu\right]dx\,dy-\int\limits_\Gamma\varphi(x,\ y)\frac{\partial z}{\partial n}ds. \tag{16}$$

Так как последнее слагаемое в формуле (16) не зависит от функции u, то краевая задача (1)—(2) эквивалентна вариационной задаче для функционала

$$\Phi[u]=\iint\limits_G\left[\left(\frac{\partial u}{\partial x}\right)^2+\left(\frac{\partial u}{\partial y}\right)^2-2fu\right]dx\,dy \tag{17}$$

в классе функций K_1.

В частном случае, если $f=f(x,\ y)\equiv 0$, то получаем уравнение Лапласа $\Delta u=0$, причем краевая задача (1)—(2) есть известная *задача Дирихле*. Решением этой задачи, как вытекает из формулы (17), является функция u из класса K_1, минимизирующая *интеграл Дирихле*

$$\Phi[u]=\iint\limits_G\left[\left(\frac{\partial u}{\partial x}\right)^2+\left(\frac{\partial u}{\partial y}\right)^2\right]dx\,dy.$$

§ 6. Идея метода Ритца

Метод Ритца служит для приближенного решения вариационной задачи.

Для простоты рассмотрим этот метод для функционала вида

$$F[u]=(Lu,\ u)-2(f,\ u), \tag{1}$$

определенного на некотором линейном множестве $K=\{u\}$, где L — положительный линейный оператор и f — заданная непрерывная функция. Предполагается, что функции класса K удовлетворяют линейным краевым условиям

$$R[u]=\varphi(P), \tag{2}$$

где R — известный линейный функционал и φ — заданная постоянная величина или функция.

Построим последовательность достаточно гладких линейно независимых функций $u_0(P)$, $u_1(P)$, ..., $u_n(P)$, где $u_0(P)$ удовлетворяет неоднородным краевым условиям

$$R[u_0]=\varphi(P), \tag{3}$$

а $u_i(P)$ $(i=1,\ 2,\ \ldots,\ n)$ — однородным краевым условиям

$$R[u_i]=0 \quad (i=1,\ 2,\ \ldots,\ n). \tag{4}$$

Составим линейную комбинацию

$$u(P;\ c_1,\ c_2,\ \ldots,\ c_n)=u_0(P)+\sum_{i=1}^{n} c_i u_i(P). \tag{5}$$

Так как $R[u]=R[u_0]+\sum\limits_{i=1}^{n} c_i\cdot 0=\varphi(P)$, то $u\in K$ при любых постоянных c_1, c_2, ..., c_n.

Приближенное решение вариационной задачи (1)—(2) будем искать в виде (5). Для этого подставим $u(P;\ c_1,\ c_2,\ \ldots,\ c_n)$ в функционал (1). Тогда получим

$$F[u]=\Phi(c_1,\ c_2,\ \ldots,\ c_n), \tag{6}$$

где Φ — известная функция, зависящая от n переменных c_1, c_2, ..., c_n. Подберем коэффициенты c_1, c_2, ..., c_n таким образом, чтобы $F[u]$ было минимальным. Это дает систему уравнений

$$\frac{\partial\Phi}{\partial c_1}=0,\quad \frac{\partial\Phi}{\partial c_2}=0,\quad \ldots,\quad \frac{\partial\Phi}{\partial c_n}=0,$$

из которой определяются постоянные c_i $(i=1,\ 2,\ldots,n)$ в формуле (5).

Таким образом, вариационная задача (1)—(2) приближенно сводится к задаче об отыскании экстремума функции $\Phi(c_1,\ c_2,\ \ldots,\ c_n)$

многих переменных. Точность решения, вообще говоря, возрастает при увеличении числа переменных функции Ф.

В следующих параграфах мы рассмотрим применение метода Ритца к конкретным краевым задачам.

§ 7. Метод Ритца для простейшей краевой задачи

Пусть дано линейное дифференциальное уравнение

$$\frac{d}{dx}[p(x)y']+q(x)y=f(x) \tag{1}$$

с простейшими краевыми условиями

$$y(a)=A, \quad y(b)=B, \tag{2}$$

где $p(x)$, $q(x)$, $f(x)\in C[a, b]$, причем $p(x)>0$ при $a\leqslant x\leqslant b$. Согласно результатам § 4 (замечание 2°) краевая задача (1)—(2) при известных условиях эквивалентна вариационной задаче для функционала

$$F[y]=\int_a^b [p(x)y'^2-q(x)y^2+2f(x)y]\,dx \tag{3}$$

на множестве функций $y\in C^{(2)}[a, b]$, удовлетворяющих краевым условиям (2).

Для решения вариационной задачи (3)—(2) применим метод Ритца. Выберем систему линейно независимых функций (координатные функции) $u_0(x)$, $u_1(x)$, $u_2(x)$, ..., $u_n(x)$ таких, что $u_0(a)=A$, $u_0(b)=B$, а остальные функции $u_i(x)$ $(i>0)$ удовлетворяют однородным краевым условиям, т. е. $u_i(a)=u_i(b)=0$ $(i=1, 2, \ldots, n)$. Решение вариационной задачи будем искать в виде линейной комбинации

$$y(x)=u_0(x)+\sum_{i=1}^{n} c_i u_i(x), \tag{4}$$

где c_i $(i=1, 2, \ldots, n)$ — некоторые постоянные. Очевидно, функция, определенная равенством (4), удовлетворяет заданным краевым условиям, т. е. $y(a)=A$, $y(b)=B$.

Коэффициенты c_1, c_2, ..., c_n подберем так, чтобы функция $y(x)$ давала экстремум функционалу (3). Подставляя выражение (4) в формулу (3), получаем

$$F[y]=\int_a^b \left\{p(x)\left[u_0'(x)+\sum_{i=1}^{n} c_i u_i'(x)\right]^2-q(x)\left[u_0(x)+\sum_{i=1}^{n} c_i u_i(x)\right]^2+\right.$$
$$\left.+2f(x)\left[u_0(x)+\sum_{i=1}^{n} c_i u_i(x)\right]\right\}dx\equiv\psi(c_1, c_2, \ldots, c_n),$$

где $\psi(c_1, c_2, \ldots, c_n)$ — квадратичная функция переменных $c_1, c_2, \ldots, c_n$. Как известно, для того чтобы дифференцируемая функция $\psi(c_1, c_2, \ldots, c_n)$ при некоторых значениях $c_1, c_2, \ldots, c_n$ имела экстремум, необходимо соблюдение для этих значений следующих условий:

$$\frac{\partial\psi}{\partial c_1}=0, \quad \frac{\partial\psi}{\partial c_2}=0, \quad \ldots, \quad \frac{\partial\psi}{\partial c_n}=0. \tag{5}$$

Система (5) является линейной относительно искомых коэффициентов $c_1, c_2, \ldots, c_n$, причем число уравнений равно числу неизвестных. Составив систему (5) и решив ее, если это возможно, найдем коэффициенты c_i $(i=1, 2, \ldots, n)$, после чего решение вариационной задачи, а следовательно, и решение исходной краевой задачи дается формулой (4). В этом и состоит формальный аспект метода Ритца для краевой задачи (1)—(2). Оценка погрешности этого метода представляет собой относительно трудную задачу [3], и разбирать ее здесь не будем. Заметим только, что точность решения в большой степени зависит от удачного подбора координатных функций, и, вообще говоря, возрастает с увеличением их числа.

Пример 1. Найти решение уравнения [1]

$$y''+(1+x^2)\,y+1=0,$$

удовлетворяющее краевым условиям $y(-1)=y(1)=0$.

Решение. За систему координатных функций $\{u_i(x)\}$ принимаем полиномы, расположенные по степеням x^2, удовлетворяющие однородным краевым условиям:

$$u_0(x)=0, \quad u_1(x)=1-x^2, \quad u_2(x)=1-x^4, \quad \ldots, \quad u_n(x)=1-x^{2n}.$$

Для простоты выкладок возьмем лишь три координатные функции, т. е. будем искать функцию $y=y(x)$ в виде суммы

$$y=c_1(1-x^2)+c_2(1-x^4). \tag{6}$$

Данное уравнение, где $p(x)=1$, $q(x)=1+x^2$, $f(x)=-1$, очевидно, является самосопряженным. Составляем для него соответствующий функционал

$$F[y]=\int_{-1}^{1}[y'^2-(1+x^2)\,y^2-2y]\,dx.$$

Заменяя y его выражением (6), получаем

$$F[y]=\int_{-1}^{1}\{(2c_1x+4c_2x^3)^2-(1+x^2)\,[c_1(1-x^2)+ \\ +c_2(1-x^4)]^2-2\,[c_1(1-x^2)+c_2(1-x^4)]\}\,dx.$$

Частные производные $\frac{\partial F}{\partial c_1}$, $\frac{\partial F}{\partial c_2}$ можно найти дифференцированием

интеграла $F[y]$ по параметрам c_1 и c_2:

$$\frac{\partial F}{\partial c_1}=\int_{-1}^{1}\{4x(2c_1x+4c_2x^3)-$$
$$-(1+x^2)\cdot 2(1-x^2)[c_1(1-x^2)+c_2(1-x^4)]-2(1-x^2)\}\,dx=$$
$$=8\left(\frac{38}{105}c_1+\frac{4}{9}c_2-\frac{1}{3}\right),$$

$$\frac{\partial F}{\partial c_2}=\int_{-1}^{1}\{8x^3(2c_1x+4c_2x^3)-$$
$$-2(1+x^2)(1-x^4)[c_1(1-x^2)+c_2(1-x^4)]-2(1-x^4)\}\,dx=$$
$$=8\left(\frac{4}{9}c_1+\frac{2488}{3645}c_2-\frac{2}{5}\right).$$

Приравнивая эти производные нулю, получаем систему уравнений

$$\frac{38}{105}c_1+\frac{4}{9}c_2=\frac{1}{3},\qquad \frac{4}{9}c_1+\frac{2488}{3645}c_2=\frac{2}{5},$$

откуда находим, что $c_1=0{,}988$, $c_2=-0{,}054$. Подставляя найденные значения c_1 и c_2 в формулу (6), получаем приближенное выражение для искомого решения:

$$y=0{,}934-0{,}988x^2+0{,}054x^4. \tag{7}$$

§ 8. Приложение метода Ритца к решению краевой задачи Штурма — Лиувилля

Рассмотрим однородное самосопряженное дифференциальное уравнение

$$[p(x)y']'+[q(x)+\lambda\rho(x)]y=0 \tag{1}$$

с однородными краевыми условиями

$$\alpha_0y(a)+\alpha_1y'(a)=0,\quad \beta_0y(b)+\beta_1y'(b)=0, \tag{2}$$

где $p(x)>0$, $|\alpha_0|+|\alpha_1|\neq 0$, $|\beta_0|+|\beta_1|\neq 0$, $p(x)$, $q(x)$, $\rho(x)$ — непрерывные функции и λ — параметр.

Очевидно, функция $y\equiv 0$ есть решение дифференциального уравнения (1), удовлетворяющее краевым условиям (2). Однако обычно представляют интерес нетривиальные решения краевой задачи (1) — (2). Отыскание нетривиальных решений дифференциального уравнения (1), удовлетворяющих однородным краевым условиям (2), называется *задачей Штурма — Лиувилля*. С этой задачей часто приходится иметь дело в уравнениях математической физики. Те значения параметра λ, при которых существуют нетривиальные значения задачи (1) — (2), называются *собственными значениями* или *собственными числами* задачи Штурма — Лиувилля, а соответствующие им нетривиальные решения — *собственными функциями* или *собственными ре-*

шениями этой задачи. Ограничимся рассмотрением уравнения (1) при простейших однородных краевых условиях

$$y(a)=0, \quad y(b)=0. \tag{3}$$

Покажем, как, используя метод Ритца, можно приближенно решить соответствующую задачу Штурма — Лиувилля.

Для этого, как указано в § 6, для уравнения (1) построим соответствующий функционал

$$F[y]=\int_a^b \{p(x)\,y'^2-[q(x)+\lambda\rho(x)]\,y^2\}\,dx. \tag{4}$$

Будем искать функцию $y=y(x, \lambda)$, дающую экстремум этому функционалу и такую, что $y(a, \lambda)=0$, $y(b, \lambda)=0$. Те значения параметра λ, при которых наша вариационная задача имеет нетривиальные решения, при известных условиях (см. § 3) являются искомыми собственными значениями рассматриваемой задачи Штурма — Лиувилля.

Искомую функцию y приближенно представим в виде линейной комбинации координатных функций

$$y=\sum_{i=1}^{n} c_i u_i(x), \tag{5}$$

где $u_i(a)=u_i(b)=0$ $(i=1, 2, \ldots, n)$. Ввиду однородности задачи полагаем $u_0(x)\equiv 0$. Подставив выражение (5) в интеграл (4) и произведя соответствующие выкладки, будем иметь

$$F[y]=\psi(c_1, c_2, \ldots, c_n, \lambda),$$

где ψ — квадратичная форма (линейная однородная функция второй степени) от переменных $c_1, c_2, \ldots, c_n$.

Коэффициенты $c_1, c_2, \ldots, c_n$ находим, используя необходимые условия для экстремума функции ψ. Это дает нам линейную однородную систему

$$\frac{\partial\psi}{\partial c_1}=0, \quad \frac{\partial\psi}{\partial c_2}=0, \quad \ldots, \quad \frac{\partial\psi}{\partial c_n}=0. \tag{6}$$

Как известно, для того чтобы однородная система (6) имела нетривиальное решение, необходимо и достаточно, чтобы определитель ее $\Delta(\lambda)$, очевидно, зависящий от параметра λ, был равен нулю. Таким образом, для определения собственных значений получаем алгебраическое уравнение n-й степени

$$\Delta(\lambda)=0, \tag{7}$$

которое называется *характеристическим уравнением* или *уравнением частот* задачи Штурма — Лиувилля. Решив характеристическое уравнение (7), находим первые n собственных значений $\lambda_1, \lambda_2, \ldots, \lambda_n$.

Для определения коэффициентов $c_1, c_2, \ldots, c_n$ следует каждое из полученных собственных значений λ_i подставить в систему (6)

и найти соответствующие нетривиальные решения этой системы. Собственные функции $y=y(x, \lambda_j)$ $(j=1, 2, \ldots, n)$ определяются из формулы (5), где коэффициенты c_i $(i=1, 2, \ldots, n)$ имеют найденные выше значения.

Заметим, что методом Ритца можно отыскать, разумеется приближенно, лишь конечное число собственных значений задачи Штурма — Лиувилля (как правило, такие задачи имеют бесконечное множество собственных значений), причем чем больше используется координатных функций, тем больше, вообще говоря, находим собственных значений и выше точность вычислений.

Пример 1. Методом Ритца определить первые два собственных значения и первые две собственные функции задачи Штурма — Лиувилля для уравнения

$$y''+\lambda y=0 \tag{8}$$

при краевых условиях

$$y(0)=y(1)=0. \tag{9}$$

Решение. Функционал (4) для данного уравнения (8) имеет вид

$$F[y]=\int_0^1 (y'^2-\lambda y^2)\,dx. \tag{10}$$

Учитывая краевые условия (9), выбираем, например, следующие координатные функции:

$$u_1(x)=x(1-x), \quad u_2(x)=x^2(1-x)$$

и соответственно полагаем

$$y=c_1(x-x^2)+c_2(x^2-x^3), \tag{11}$$

где c_1 и c_2 — постоянные коэффициенты, не равные нулю одновременно $(c_1^2+c_2^2>0)$.

Подставляя выражение (11) в формулу (10), будем иметь

$$F[y]=\int_0^1 \{[c_1(1-2x)+c_2(2x-3x^2)]^2-\lambda\,[c_1(x-x^2)+c_2(x^2-x^3)]^2\}\,dx.$$

Отсюда, дифференцируя по параметрам c_1 и c_2 под знаком интеграла, получим

$$\frac{1}{2}\frac{\partial F}{\partial c_1}=\int_0^1 \{(1-2x)[c_1(1-2x)+c_2(2x-3x^2)]\}-$$
$$-\lambda(x-x^2)[c_1(x-x^2)+c_2(x^2-x^3)]\}\,dx=$$
$$=\left(\frac{1}{3}c_1+\frac{1}{6}c_2\right)-\lambda\left(\frac{1}{30}c_1+\frac{1}{60}c_2\right),$$

$$\frac{1}{2}\frac{\partial F}{\partial c_2}=\int_0^1 \{(2x-3x^2)[c_1(1-2x)+$$
$$+c_2(2x-3x^2)]-\lambda(x^2-x^3)[c_1(x-x^2)+c_2(x^2-x^3)]\}\,dx=$$
$$=\left(\frac{1}{6}c_1+\frac{2}{15}c_2\right)-\lambda\left(\frac{1}{60}c_1+\frac{1}{105}c_2\right).$$

Приравнивая нулю производные $\frac{\partial F}{\partial c_1}$ и $\frac{\partial F}{\partial c_2}$, приходим к системе

$$\left(\frac{1}{3}c_1+\frac{1}{6}c_2\right)\left(1-\frac{\lambda}{10}\right)=0,$$
$$\frac{1}{6}c_1\left(1-\frac{\lambda}{10}\right)+\frac{2}{15}c_2\left(1-\frac{\lambda}{14}\right)=0$$

или

$$\left.\begin{aligned}(2c_1+c_2)\left(1-\frac{\lambda}{10}\right)&=0,\\ 5c_1\left(1-\frac{\lambda}{10}\right)+4c_2\left(1-\frac{\lambda}{14}\right)&=0.\end{aligned}\right\} \tag{12}$$

Система (12) имеет ненулевое решение c_1, c_2 тогда и только тогда, когда определитель ее равен нулю. Приравнивая нулю определитель системы (12), получим *характеристическое уравнение*

$$\begin{vmatrix} 2\left(1-\frac{\lambda}{10}\right) & \left(1-\frac{\lambda}{10}\right)\\ 5\left(1-\frac{\lambda}{10}\right) & 4\left(1-\frac{\lambda}{14}\right)\end{vmatrix}=0$$

или, после упрощений,

$$\left(1-\frac{\lambda}{10}\right)\left(3-\frac{\lambda}{14}\right)=0. \tag{13}$$

Отсюда находим приближенные собственные значения задачи

$$\lambda_1=10,\quad \lambda_2=42. \tag{14}$$

Коэффициенты c_1 и c_2 определяем из системы (12). При $\lambda=\lambda_1=10$ имеем $c_1=c$, $c_2=0$. Следовательно, первая собственная функция нашей краевой задачи в силу формулы (11) есть

$$y_1\approx c(x-x^2)\quad (c\neq 0).$$

Полагая $\lambda=\lambda_2=42$ в системе (12), будем иметь

$$2c_1+c_2=0,\quad 16c_1-8c_2=0;$$

отсюда $c_1=c$, $c_2=-2c$. Подставляя последние выражения в формулу (11), получаем вторую собственную функцию

$$y_2\approx c(x-3x^2+2x^3)\quad (c\neq 0).$$

В данном случае известно точное решение краевой задачи (8)—(9). А именно, собственные значения имеют вид $\lambda_n=n^2\pi^2$ $(n=1, 2, \ldots)$, а соответствующие собственные функции определяются формулой

$$y_n=c\sin n\pi x\quad (n=1, 2, \ldots),$$

где $c\neq 0$. В частности, получаем $\lambda_1=\pi^2=9{,}87$ и $\lambda_2=4\pi^2=39{,}48$. Таким образом, из приближенных собственных значений (14) первое определено примерно с относительной погрешностью 1,3%, а второе—с относительной погрешностью 6,4%.

§ 9. Метод Ритца для задачи Дирихле

Будем искать решение уравнения Лапласа

$$\Delta u = 0 \text{ при } (x, y) \in G \tag{1}$$

и

$$u|_\Gamma = f(x, y), \tag{2}$$

где Γ — простой замкнутый контур, ограничивающий конечную область G, а функция $f(x, y)$ непрерывна на Γ. Согласно § 6 эта краевая задача эквивалентна вариационной задаче для функционала

$$F[y] = \iint\limits_G \left\{ \left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 \right\} dx\, dy \tag{3}$$

в классе функций, имеющих непрерывные частные производные до второго порядка включительно в замкнутой области $G+\Gamma$ и удовлетворяющих на границе Γ краевому условию (2). Построим конечную систему линейно независимых функций (координатные функции)

$$u_0(x, y),\ u_1(x, y),\ \ldots,\ u_n(x, y) \in C^{(2)}(G+\Gamma)$$

таких, что

$$u_0(x, y)|_\Gamma = f(x, y),\ u_i(x, y)|_\Gamma = 0 \quad (i = 1, 2, \ldots, n).$$

Тогда линейная комбинация

$$\varphi(x, y) = u_0(x, y) + \sum_{i=1}^{n} c_i u_i(x, y) \tag{4}$$

принадлежит классу допустимых функций при любых постоянных $c_1, c_2, \ldots, c_n$. Формулу (4) можно записать короче:

$$\varphi(x, y) = \sum_{i=0}^{n} c_i u_i(x, y), \tag{4'}$$

где $c_0 = 1$. Подставляя выражение (4′) в функционал (3), получим

$$F[\varphi] = \iint\limits_G \left[\left(\sum_{i=0}^{n} c_i \frac{\partial u_i}{\partial x} \right)^2 + \left(\sum_{i=0}^{n} c_i \frac{\partial u_i}{\partial y} \right)^2 \right] dx\, dy. \tag{5}$$

Подберем коэффициенты $c_1, c_2, \ldots, c_n$ так, чтобы функция $F[\varphi]$ имела минимум. Для этого необходимо выполнение условий

$$\frac{\partial}{\partial c_j} F[\varphi] = 2 \iint\limits_G \left[\left(\sum_{i=0}^{n} c_i \frac{\partial u_i}{\partial x} \right) \frac{\partial u_j}{\partial x} + \left(\sum_{i=0}^{n} c_i \frac{\partial u_i}{\partial y} \right) \frac{\partial u_j}{\partial y} \right] dx\, dy = 0 \quad (6)$$

$$(j = 1, 2, \ldots, n)$$

или

$$\left.\begin{aligned} &[u_0, u_1] + c_1 [u_1, u_1] + \ldots + c_n [u_n, u_1] = 0, \\ &[u_0, u_2] + c_1 [u_1, u_2] + \ldots + c_n [u_n, u_2] = 0, \\ &[u_0, u_n] + c_1 [u_1, u_n] + \ldots + c_n [u_n, u_n] = 0, \end{aligned}\right\} \quad (6')$$

где

$$[u_i, u_j] = \iint\limits_G \left(\frac{\partial u_i}{\partial x} \frac{\partial u_j}{\partial x} + \frac{\partial u_i}{\partial y} \frac{\partial u_j}{\partial y} \right) dx\, dy,$$

причем

$$[u_i, u_j] = [u_j, u_i]. \quad (7)$$

Из линейной системы (6′) определяются коэффициенты $c_1, c_2, \ldots, c_n$.

Функция $\varphi(x, y)$ с коэффициентами, определенными из системы (6′), представляет собой приближенное решение задачи Дирихле. Точность приближения зависит от выбора координатных функций $u_k(x, y)$ и от числа этих функций [3].

Для ознакомления с применением метода Ритца к более общим краевым задачам для уравнений с частными производными следует обратиться к более подробным руководствам [1], [3].

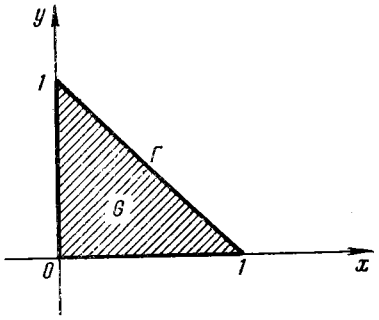

Рис. 82.

Пример 1. Найти функцию $u = u(x, y)$, гармоническую в области G: $x > 0$, $y > 0$, $x + y < 1$ и удовлетворяющую на границе Г: $x = 0$, $y = 0$, $x + y = 1$ (рис. 82) условию

$$u|_\Gamma = x^2 + y^2. \quad (8)$$

Решение. Выберем следующую систему координатных функций:

$$u_0(x, y) = x^2 + y^2, \quad u_1(x, y) = xy(1 - x - y),$$
$$u_2(x, y) = x^2 y(1 - x - y), \quad u_3(x, y) = xy^2(1 - x - y), \ldots$$

и составим линейную комбинацию

$$\varphi(x, y) = x^2 + y^2 + c_1 xy(1 - x - y) + c_2 x^2 y(1 - x - y) + c_3 xy^2(1 - x - y). \quad (9)$$

Легко проверить, что функция $\varphi(x, y)$ удовлетворяет краевому условию (8) при любых значениях постоянных c_1, c_2, c_3.

Для составления системы (6′) подсчитываем коэффициенты при неизвестных c_1, c_2, c_3 и свободные члены.

$$[u_0, u_1]=[u_1, u_0]=\iint\limits_G [2x(y-2xy-y^2)+2y(x-x^2-2xy)]\,dx\,dy,$$

$$[u_0, u_2]=[u_2, u_0]=\iint\limits_G [2x(2xy-3x^2y-2xy^2)+2y(x^2-x^3-2x^2y)]\,dx\,dy,$$

$$[u_0, u_3]=[u_3, u_0]=\iint\limits_G [2x(y^2-2xy^2-y^3)+2y(2xy-2x^2y-3xy^2)]\,dx\,dy,$$

$$[u_1, u_1]=\iint\limits_G [(y-2xy-y^2)^2+(x-x^2-2xy)^2]\,dx\,dy,$$

$$[u_1, u_2]=[u_2, u_1]=\iint\limits_G [(y-2xy-y^2)(2xy-3x^2y-2xy^2)+$$
$$+(x-x^2-2xy)(x^2-x^3-2x^2y)]\,dx\,dy,$$

$$[u_1, u_3]=[u_3, u_1]=\iint\limits_G [(y-2xy-y^2)(y^2-2xy^2-y^3)+$$
$$+(x-x^2-2xy)(2xy-2x^2y-3xy^2)]\,dx\,dy,$$

$$[u_2, u_2]=\iint\limits_G (2xy-3x^2y-2xy^2)^2+(x^2-x^3-2x^2y)^2]dx\,dy,$$

$$[u_2, u_3]=[u_3, u_2]=\iint\limits_G [(2xy-3x^2y-2xy^2)(y^2-2xy^2-y^3)+$$
$$+(x^2-x^3-2x^2y)(2xy-2x^2y-3xy^2)]\,dx\,dy,$$

$$[u_3, u_3]=\iint\limits_G [(y^2-2xy^2-y^3)^2+(2xy-2x^2y-3xy^2)^2]\,dx\,dy.$$

Результаты вычислений приведены в таблице 74.

Таблица 74

Значения коэффициентов при неизвестных

j	$[u_0, u_j]$	$[u_1, u_j]$	$[u_2, u_j]$	$[u_3, u_j]$
1	$-\frac{1}{30}$	$\frac{1}{90}$	$\frac{1}{252}$	$\frac{1}{252}$
2	$-\frac{1}{90}$	$\frac{1}{252}$	$\frac{3}{1120}$	$\frac{1}{70}$
3	$-\frac{1}{90}$	$\frac{1}{252}$	$\frac{1}{70}$	$\frac{3}{1120}$

Отсюда линейная система для определения коэффициентов запишется в виде

$$\left.\begin{aligned} &\frac{1}{90}c_1+\frac{1}{252}c_2+\frac{1}{252}c_3=\frac{1}{30},\\ &\frac{1}{252}c_1+\frac{3}{1120}c_2+\frac{1}{70}c_3=\frac{1}{90},\\ &\frac{1}{252}c_1+\frac{1}{70}c_2+\frac{3}{1120}c_3=\frac{1}{90}. \end{aligned}\right\} \tag{10}$$

Решив систему (10), находим

$$c_1=\frac{3031}{997}\approx 3{,}0401; \qquad c_2=c_3=-\frac{56}{997}=-0{,}0562.$$

Подставляя найденные значения величин c_1, c_2, c_3 в формулу (9), получаем приближенное решение нашей задачи:

$$\varphi(x,\ y)=x^2+y^2+xy\,(1-x-y)\,[3{,}0401-0{,}0562\,(x+y)].$$

ЛИТЕРАТУРА К ГЛАВЕ VI

[1] Коллатц Л., Численные методы решения дифференциальных уравнений, ИЛ, 1953, гл. III и IV.
[2] Березин И. А. и Жидков Н. П., Методы вычислений, т. 2, Физматгиз, 1959, гл. X.
[3] Михлин С. Г., Вариационные методы в математической физике, Гостехиздат, 1957, гл. III.

ГЛАВА VII

ИНТЕГРАЛЬНЫЕ УРАВНЕНИЯ

§ 1. Основные виды линейных интегральных уравнений

Под *интегральным уравнением* понимается уравнение, содержащее неизвестную функцию $y(x)$ под знаком определенного интеграла [1]—[15]. В дальнейшем мы ограничимся рассмотрением *линейных интегральных уравнений*, в которые неизвестная функция входит лишь в первой степени (линейно).

Приведем некоторые наиболее часто встречающиеся типы интегральных уравнений. Уравнение вида

$$\int_a^b K(x, s)\, y(s)\, ds = f(x), \tag{1}$$

где $K(x, s)$ (*ядро*) и $f(x)$ — известные функции, называется *интегральным уравнением Фредгольма первого рода.*

Уравнение вида

$$y(x) - \lambda \int_a^b K(x, s)\, y(s)\, ds = f(x), \tag{2}$$

где λ — числовой параметр, носит название *интегрального уравнения Фредгольма второго рода.*

Параметр λ вводится по следующим соображениям: при данном значении λ интегральное уравнение (2) не всегда имеет решения. Варьируя параметр λ, можно добиться того, чтобы решение уравнения (2) существовало. Параметр λ можно также ввести в левую часть уравнения Фредгольма первого рода (1).

Если в (2) $f(x) \equiv 0$, то получается однородное уравнение

$$y(x) = \lambda \int_a^b K(x, s)\, y(s)\, ds, \tag{3}$$

допускающее нулевое (тривиальное) решение $y \equiv 0$. Те значения параметра λ, при которых однородное интегральное уравнение (3) имеет

нетривиальные решения, называются *собственными значениями* (*собственными числами*) ядра $K(x, s)$ или соответствующего уравнения (2), а отвечающие им ненулевые решения — *собственными функциями*. Основной результат теории следующий (теорема Фредгольма) [10], [11]: 1) если λ не есть собственное значение ядра $K(x, s)$, то соответствующее неоднородное интегральное уравнение Фредгольма (2) с регулярным ядром $K(x, s)$ и непрерывным свободным членом $f(x)$ имеет единственное непрерывное решение $y(x)$ $(a \leqslant x \leqslant b)$; 2) если же λ есть собственное значение, то уравнение (2) или не имеет решений, или же допускает бесчисленное множество их.

В приложениях важную роль играют уравнения Фредгольма второго рода с *симметрическим ядром* $K(x, s)$, т. е. таким, что

$$K(x, s) = K(s, x).$$

Симметрическое ядро обладает следующими свойствами (см., например, [5], [11]):

1) для всякого симметрического ядра существует по меньшей мере одно собственное значение;

2) все собственные значения симметрического ядра действительны;

3) собственные функции $\varphi(x)$ и $\psi(x)$ симметрического ядра, соответствующие различным собственным значениям λ и μ $(\lambda \neq \mu)$, ортогональны между собой на основном промежутке (a, b), т. е.

$$\int_a^b \varphi(x)\,\psi(x)\,dx = 0.$$

Пример 1. Пусть простой замкнутый кусочно-гладкий контур

$$x = \xi(t),\quad y = \eta(t)\quad (0 \leqslant t \leqslant T) \qquad (\Gamma)$$

ограничивает конечную область G. Тогда функция $u(x, y)$, дающая решение соответствующей задачи Дирихле (гл. V, § 4), т. е. такая, что

$$\Delta u \equiv \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 \quad \text{при } (x, y) \in G$$

и

$$u = f(t) \text{ при } (x, y) \in \Gamma$$

($f(t)$ — известная функция), может быть представлена в виде (см. [11], [16])

$$u(x, y) = \oint_\Gamma \mu(t)\,\frac{\partial\theta}{\partial t}\,dt,$$

где

$$\theta(t, x, y) = \operatorname{arctg}\frac{\eta(t) - y}{\xi(t) - x}$$

и функция $\mu(t)$ удовлетворяет интегральному уравнению

$$\pi\mu(s) + \int_0^T K(s, t)\,\mu(t)\,dt = f(s),$$

ядро которого есть

$$K(s, t)=\frac{\partial}{\partial t}\operatorname{arctg}\frac{\eta(t)-\eta(s)}{\xi(t)-\xi(s)}.$$

В частности, например, для эллипса

$$x=a\cos t, \quad y=b\sin t \quad (b\leqslant a)$$

будем иметь

$$K(s, t)=\frac{ab}{(a^2+b^2)-(a^2-b^2)\cos(s+t)}.$$

В приложениях встречаются также интегральные уравнения вида

$$\int_a^x K(x, s)\, y(s)\, ds=f(x) \tag{4}$$

и

$$y(x)-\lambda\int_a^x K(x, s)\, y(s)\, ds=f(x), \tag{5}$$

которые носят названия *интегральных уравнений Вольтерра* соответственно первого и второго рода. Вводя функцию

$$K^*(x, s)=\begin{cases} K(x, s) & \text{при } a\leqslant s\leqslant x, \\ 0 & \text{при } s>x, \end{cases}$$

уравнения Вольтерра (4) и (5) можно записать в виде соответствующих уравнений Фредгольма с ядром $K^*(x, s)$. Таким образом, теория уравнений Вольтерра сводится к теории уравнений Фредгольма; однако в некоторых случаях уравнения Вольтерра полезно изучать независимо.

Примером уравнения Вольтерра первого рода является *обобщенное уравнение Абеля*

$$\int_0^x \frac{y(s)\, ds}{(x-s)^\alpha}=f(x) \quad (0<\alpha<1), \tag{6}$$

где $f(x)$ — известная непрерывно дифференцируемая функция. Решение уравнения (6) дается формулой (см. [2], [11])

$$y(x)=\frac{\sin\alpha\pi}{\pi}\left[\frac{f(0)}{x^{1-\alpha}}+\int_0^x \frac{f'(s)\, ds}{(x-s)^{1-\alpha}}\right],$$

в чем можно убедиться непосредственно.

Заметим, что если ядро $K(x, s)$ и $f(x)$ — непрерывно дифференцируемые функции, причем $K(x, x)\neq 0$ при $a\leqslant x\leqslant b$, то уравнение Вольтерра первого рода (4) сводится к уравнению Вольтерра

второго рода (5). Действительно, дифференцируя уравнение (4) по x, будем иметь

$$K(x, x)\, y(x) + \int_a^x K'_x(x, s)\, y(s)\, ds = f'(x);$$

отсюда

$$y(x) + \int_a^x K_1(x, s)\, y(s)\, ds = f_1(x) \qquad (a \leqslant x \leqslant b),$$

где

$$K_1(x, s) = \frac{K'_x(x, s)}{K(x, x)}, \quad f_1(x) = \frac{f'(x)}{K(x, x)}.$$

Поэтому в дальнейшем мы не будем отдельно заниматься интегральными уравнениями Вольтерра первого рода.

К линейным интегральным уравнениям может быть приведено большое количество задач математической физики.

В основном мы будем заниматься интегральным уравнением Фредгольма второго рода и частично уравнением Вольтерра второго рода (для краткости в дальнейшем мы их будем именовать просто уравнениями Фредгольма и Вольтерра).

Основными проблемами здесь являются следующие:

1) нахождение приближенного или точного решения неоднородного интегрального уравнения при заданном значении параметра λ;

2) нахождение собственных значений и соответствующих собственных функций однородного интегрального уравнения.

§ 2. Связь между дифференциальными уравнениями и уравнениями Вольтерра

Рассмотрим линейное дифференциальное уравнение

$$\frac{d^2u}{dx^2} + p(x)\frac{du}{dx} + q(x)\, u = f(x) \qquad (a \leqslant x \leqslant b) \tag{1}$$

при начальных условиях

$$u(a) = A, \qquad u'(a) = B. \tag{2}$$

Полагая

$$\frac{d^2u}{dx^2} = y(x), \tag{3}$$

после двукратного интегрирования получим

$$\frac{du}{dx} = \int_a^x y(s)\, ds + C_1$$

и

$$u=\int_a^x ds\int_a^s y(t)\,dt+C_1(x-a)+C_2.$$

Изменяя порядок интегрирования в двойном интеграле, будем иметь

$$\int_a^x ds\int_a^s y(t)\,dt=\int_a^x dt\int_t^x y(t)\,ds=\int_a^x (x-t)\,y(t)\,dt\equiv\int_a^x (x-s)\,y(s)\,ds.$$

Кроме того, из начальных условий (2) при $x=a$ находим $C_1=B$, $C_2=A$. Поэтому

$$\frac{du}{dx}=\int_a^x y(s)\,ds+B \tag{4}$$

и

$$u(x)=\int_a^x (x-s)\,y(s)\,ds+B(x-a)+A. \tag{5}$$

Подставляя выражения (3), (4) и (5) в дифференциальное уравнение (1), будем иметь интегральное уравнение Вольтерра

$$y(x)+\int_a^x K(x,s)\,y(s)\,ds=F(x), \tag{6}$$

где

$$K(x,s)=p(x)+q(x)(x-s),$$

$$F(x)=f(x)-Bp(x)-[B(x-a)+A]\,q(x).$$

Зная функцию $y(x)$, можно по формуле (5) найти решение $u(x)$ и производную $u'(x)$; таким образом, интегральное уравнение (6) включает в себя всю информацию, связанную с начальной задачей (задачей Коши) для дифференциального уравнения (1).

Аналогичный результат получается для линейного дифференциального уравнения n-го порядка [5].

Обратно, если ядро $K(x,s)$ есть целый полином относительно s степени n, т. е.

$$K(x,s)=\sum_{m=0}^{n} a_m(x)\,s^m,$$

то путем последовательного дифференцирования интегрального уравнения (6) мы придем к задаче Коши для некоторого линейного дифференциального уравнения.

Пример 1. Решить интегральное уравнение

$$y(x)+\int_0^x (2+x-s)\,y(s)\,ds=x^2. \tag{7}$$

Последовательно продифференцировав два раза, будем иметь

$$y'(x)+2y(x)+\int_0^x y(s)\,ds=2x, \tag{8}$$

$$y''(x)+2y'(x)+y(x)=2. \tag{9}$$

Из уравнений (7) и (8) при $x=0$ получаем начальные условия

$$y(0)=0, \qquad y'(0)=0.$$

Решая обычным приемом дифференциальное уравнение 9), находим

$$y(x)=2-2e^{-x}(1+x).$$

§ 3. Связь линейной краевой задачи с интегральным уравнением Фредгольма

Рассмотрим для самосопряженного дифференциального уравнения второго порядка

$$L[y]\equiv p(x)\,y''+p'(x)\,y'+q(x)\,y=f(x), \tag{1}$$

где $p(x)$, $p'(x)$, $q(x)$, $f(x)$ непрерывны на $[a, b]$ и $p(x)>0$ при $a\leqslant x\leqslant b$, однородную краевую задачу (гл. IV, § 3)

$$\left.\begin{aligned}\Gamma_a[y]&\equiv\alpha_0 y(a)+\alpha_1 y'(a)=0,\\ \Gamma_b[y]&\equiv\beta_0 y(b)+\beta_1 y'(b)=0\end{aligned}\right\} \tag{2}$$
$$(|\alpha_0|+|\alpha_1|\neq 0,\ |\beta_0|+|\beta_1|\neq 0).$$

Определение. Функция $G(x, s)$ называется *функцией Грина (функцией влияния)* [4], [11] краевой задачи (1)—(2), если выполнены следующие условия:

1) $G(x, s)$ определена и непрерывна в области $a\leqslant x\leqslant b$, $a\leqslant s\leqslant b$;
2) $L_x[G(x, s)]=0$ при $x\neq s$;
3) $\Gamma_a[G(a, s)]=0$, $\Gamma_b[G(b, s)]=0$;
4) $G'_x(s+0, s)-G'_x(s-0, s)=-\dfrac{1}{p(s)}$.

Если функция Грина $G(x, s)$ найдена, то решение краевой задачи (1)—(2) дается формулой

$$y(x)=-\int_a^b G(x, s)\,f(s)\,ds, \tag{3}$$

что нетрудно проверить непосредственно. Можно доказать, что функция Грина $G(x, s)$ симметрическая: $G(x, s)=G(s, x)$.

Рассмотрим теперь краевую задачу Штурма — Лиувилля (гл. VI, §8)

$$[p(x)y']' + [q(x) + \lambda\rho(x)]y = 0, \tag{4}$$

$$\Gamma_a[y] = \Gamma_b[y] = 0 \quad (p(x) > 0, \quad \rho(x) > 0). \tag{5}$$

Интерпретируя $\lambda\rho(x)y$ как свободный член, на основании формулы (3) непосредственно приходим к однородному интегральному уравнению Фредгольма

$$y(x) = \lambda \int_a^b K(x, s)\, y(s)\, ds, \tag{6}$$

где

$$K(x, s) = G(x, s)\,\rho(s). \tag{7}$$

Собственные значения λ задачи Штурма — Лиувилля, очевидно, будут являться собственными значениями интегрального уравнения (6).

Уравнение (6) можно свести к интегральному уравнению с симметрическим ядром. Действительно полагая $z(x) = y(x)\sqrt{\rho(x)}$ в силу формул (6) и (7), будем иметь

$$z(x) = \lambda \int_a^b K^*(x, s)\, z(s)\, ds,$$

где ядро $K^*(x, s) = G(x, s)\sqrt{\rho(x)}\sqrt{\rho(s)}$ симметрично. Следовательно, все собственные значения λ действительны.

§ 4. Метод последовательных приближений

Рассмотрим уравнение Фредгольма

$$y(x) = f(x) + \lambda \int_a^b K(x, s)\, f(s)\, ds, \tag{1}$$

где $f(x)$ и $K(x, s)$ непрерывны.

Будем искать решение в форме степенного ряда

$$y(x) = \sum_{n=0}^{\infty} \lambda^n \varphi_n(x). \tag{2}$$

Подставляя выражение (2) в интегральное уравнение (1) и приравнивая коэффициенты при одинаковых степенях параметра λ, будем иметь

$$\left.\begin{aligned} &\varphi_0(x) = f(x), \\ &\varphi_n(x) = \int_a^b K(x, s)\, \varphi_{n-1}(s)\, ds \quad (n = 1, 2, \ldots). \end{aligned}\right\} \tag{3}$$

Пусть $|K(x, s)| \leqslant M$ и $|f(x)| \leqslant N$ в области $R\{a \leqslant x \leqslant b,\ a \leqslant s \leqslant b\}$.

Из формул (3) по индукции получаем $|\varphi_n(x)| \leqslant M^n N (b-a)^n$. Поэтому сходимость ряда (2) будет обеспечена, если

$$|\lambda| < \frac{1}{M(b-a)}.$$

Приняв

$$y(x) \approx y_n(x) = \sum_{k=0}^{n} \lambda^k \varphi_k(x),$$

мы получим приближенное решение интегрального уравнения (1) с погрешностью

$$\varepsilon_n = |y(x) - y_n(x)| \leqslant \sum_{k=n+1}^{\infty} |\lambda|^k |\varphi_k(x)| \leqslant$$

$$\leqslant \sum_{k=n+1}^{\infty} N[M(b-a)|\lambda|]^k = \frac{N[M(b-a)|\lambda|]^{n+1}}{1-M(b-a)|\lambda|}. \qquad (4)$$

Формула (2) дает аналитическое относительно λ решение уравнения Фредгольма (1) в окрестности точки $\lambda = 0$. Из формул (3) вытекает, что решение (2) можно записать в виде

$$y(x) = f(x) + \lambda \sum_{n=1}^{\infty} \lambda^{n-1} \int_a^b K_n(x, s) f(s)\, ds$$

или

$$y(x) = f(x) + \lambda \int_a^b R(x, s, \lambda) f(s)\, ds, \qquad (5)$$

где

$$R(x, s, \lambda) = \sum_{n=1}^{\infty} \lambda^{n-1} K_n(x, s). \qquad (6)$$

Коэффициенты $K_n(x, s)$, так называемые *итерированные ядра*, могут быть последовательно найдены по формулам [11]

$$K_1(x, s) = K(x, s),$$

$$K_n(x, s) = \int_a^b K(x, t) K_{n-1}(t, s)\, dt \quad (n = 2, 3, \ldots).$$

Функция $R(x, s, \lambda)$ называется *резольвентой* уравнения (1) и при малых $|\lambda|$ определяется степенным рядом (6). Пользуясь аналитическим продолжением [17], резольвенту $R(x, s, \lambda)$ можно продолжить на всю комплексную плоскость параметра λ, за исключением собственных значений $\lambda_1, \lambda_2, \ldots$ (*особые точки*), которые являются

полюсами резольвенты. Тогда формула (5) дает решение интегрального уравнения (1) при любом $\lambda \neq \lambda_k$ $(k=1, 2, \ldots)$.

Рассмотрим теперь соответствующее уравнение Вольтерра

$$y(x) = f(x) + \lambda \int_a^x K(x, s) f(s)\, ds, \tag{7}$$

где $a \leqslant x \leqslant b$. Полагая

$$y(x) = \sum_{n=0}^{\infty} \lambda^n \psi_n(x), \tag{8}$$

аналогично предыдущему получим

$$\psi_0(x) = f(x), \quad \psi_n(x) = \int_a^x K(x, s)\, \psi_{n-1}(s)\, ds \quad (n=1, 2, \ldots).$$

Отсюда

$$|\psi_n(x)| \leqslant \frac{M^n N (b-a)^n}{n!} \quad (n=0, 1, 2, \ldots), \tag{9}$$

где

$$|K(x, s)| \leqslant M \text{ при } a \leqslant x \leqslant b, \quad a \leqslant s \leqslant b$$

и

$$|f(x)| \leqslant N \quad \text{при} \quad a \leqslant x \leqslant b.$$

Следовательно, ряд (8) сходится при любом λ и дает единственное решение уравнения (7). Погрешность приближенного решения

$$Y_n(x) = \sum_{k=0}^{n} \lambda^k \psi_k(x)$$

на основании оценок (9) дается формулой

$$\varepsilon_n = |y(x) - Y_n(x)| \leqslant \sum_{k=n+1}^{\infty} \frac{|\lambda|^k M^k N (b-a)^k}{k!} =$$

$$= N\left\{ e^{M|\lambda|(b-a)} - \sum_{k=0}^{n} \frac{[|\lambda| M (b-a)]^k}{k!} \right\}.$$

Пример 1. Методом последовательных приближений найти приближенное решение уравнения

$$y(x) = x + \lambda \int_0^1 \frac{y(s)}{10+x+s}\, ds \quad (0 \leqslant x \leqslant 1).$$

Полагая

$$y(x) = \sum \lambda^n \varphi_n(x), \tag{10}$$

имеем

$$\varphi_0(x)=x, \quad \varphi_1(x)=\int_0^1 \frac{s}{10+x+s}\,ds=1-(10+x)\ln\frac{11+x}{10+x}.$$

Таким образом, в качестве первого приближения можно взять

$$y_1(x)=x+\lambda\left[1-(10+x)\ln\left(1+\frac{1}{10+x}\right)\right].$$

Здесь

$$M=\max\frac{1}{10+x+s}=0{,}1, \quad N=\max_{0\leqslant x\leqslant 1} x=1.$$

Следовательно, ряд (10) сходится при $|\lambda|<\frac{1}{0{,}1\cdot(1-0)}=10$.

В частности, при $\lambda=1$ точность решения на основании (4) есть

$$|y(x)-y_1(x)|<\frac{1\cdot(0{,}1\cdot 1)^2}{1-0{,}1\cdot 1}\approx 0{,}01.$$

Заметим, что неудобством метода последовательных приближений является необходимость вычисления квадратур. Если интегралы не вычисляются точно, то приходится прибегать к численным квадратурным формулам.

§ 5. Решение интегрального уравнения методом конечных сумм

Метод основывается на приближенном вычислении определенного интеграла с помощью некоторой квадратурной формулы

$$\int_a^b F(x)\,dx=\sum_{i=1}^{n} A_i F(x_i)+R[F], \tag{1}$$

где x_i $(i=1, 2, \ldots, n)$ — абсциссы точек отрезка $[a, b]$, A_i $(i=1, 2, \ldots, n)$ — числовые коэффициенты, не зависящие от выбора функции $F(x)$, и $R[F]$ — остаточный член (ошибка) формулы (1). Обычно $A_i\geqslant 0$ и $\sum_{i=1}^{n} A_i=b-a$.

Например, в случае равноотстоящих точек $x_i=a+(i-1)h$ $(i=1, 2, \ldots, n)$, где $h=(b-a)/(n-1)$, будем иметь (см. [18]):

1) для формулы прямоугольников:

$$A_i=h \qquad (i=1, 2, \ldots, n-1), \quad A_n=0;$$

2) для общей формулы трапеций:

$$A_1=A_n=\frac{h}{2}, \quad A_2=A_3=\ldots=A_{n-1}=h;$$

3) для общей формулы Симпсона при $n=2m+1$:

$$A_1=A_{2m+1}=\frac{h}{3},$$

$$A_2=A_4=\ldots=A_{2m}=\frac{4h}{3},$$

$$A_3=A_5=\ldots=A_{2m-1}=\frac{2h}{3}.$$

Другие квадратурные формулы см. в [16], [18], [19].

Пусть теперь дано интегральное уравнение Фредгольма второго рода

$$y(x)-\lambda\int_a^b K(x,\ s)\,y(s)\,ds=f(x)\quad(a\leqslant x\leqslant b). \tag{2}$$

Выбирая точки $x_i\in[a,\ b]$ и вводя обозначения:

$$y(x_i)=y_i,\quad K(x_i,\ x_j)=K_{ij},\quad f(x_i)=f_i\quad(i,\ j=1,\ 2,\ \ldots,\ n),$$

на основании формулы (1) будем иметь

$$y_i-\lambda\sum_{j=1}^{n}A_jK_{ij}y_j=f_i+R_i,\qquad(i=1,\ 2,\ \ldots,\ n), \tag{3}$$

где R_i — соответствующие ошибки. Отбрасывая в системе (3) величины R_i, для приближенных значений Y_i решения $y(x)$ в узлах x_i $(i=1,\ 2,\ \ldots,\ n)$ получим линейную алгебраическую систему

$$Y_i-\lambda\sum_{j=1}^{n}A_jK_{ij}Y_j=f_i\qquad(i=1,\ 2,\ \ldots,\ n). \tag{4}$$

Вводя символ Кронекера

$$\delta_{ij}=\begin{cases}0, & \text{если } i\neq j,\\ 1, & \text{если } i=j,\end{cases}$$

и учитывая, что

$$Y_i=\sum_{j=1}^{n}\delta_{ij}Y_j,$$

систему (4) можем записать в виде

$$\sum_{j=1}^{n}(\delta_{ij}-\lambda A_jK_{ij})\,Y_j=f_i\qquad(i=1,\ 2,\ \ldots,\ n). \tag{4'}$$

Если

$$\Delta(\lambda)=\det(\delta_{ij}-\lambda A_jK_{ij})\neq 0, \tag{5}$$

то система (4′) имеет единственное решение Y_i, которое можно найти методом Гаусса или другими методами, разработанными для решения систем алгебраических линейных уравнений (см. например [18]).

Найдя Y_i $(i=1, 2, \ldots, n)$, для решения $y(x)$ получаем из уравнения (2) приближенное аналитическое выражение

$$Y(x)=f(x)+\lambda\sum_{j=1}^{n}A_jK(x, x_j)Y_j. \tag{6}$$

Различные между собой корни $\tilde{\lambda}_1, \tilde{\lambda}_2, \ldots, \tilde{\lambda}_m$ $(m\leqslant n)$ алгебраического уравнения $\Delta(\lambda)=0$ представляют собой, вообще говоря, приближения собственных значений ядра $K(x, s)$. Если $\tilde{Y}_{ik}^l$ $(i=1, 2, \ldots, n;\ k=1, 2, \ldots, m;\ l=1, \ldots, p_k)$ — соответствующие ненулевые решения однородной системы

$$\sum_{j=1}^{n}(\delta_{ij}-\tilde{\lambda}_kA_jK_{ij})\tilde{Y}_{jk}^l=0 \quad (i=1, 2, \ldots, n), \tag{7}$$

то собственные функции ядра приближенно определяются формулами

$$\tilde{\varphi}_{kl}(x)=\tilde{\lambda}_k\sum_{j=1}^{n}A_jK(x, x_j)\tilde{Y}_{jk}^l \quad (k=1, 2, \ldots, m;\ l=1, \ldots, p_k).$$

Оценка погрешности метода конечных сумм приведена в [16] и [19]. Заметим, что этот метод дает хорошие результаты, если ядро $K(x, s)$ и правая часть $f(x)$ достаточно гладкие функции. В противном случае полезно предварительно преобразовать соответствующим образом интегральное уравнение (см. [19]).

Метод конечных сумм может быть применен также к интегральному уравнению Фредгольма первого рода

$$\lambda\int_a^b K(x, s)\,y(s)\,ds=f(x).$$

В этом случае приближенные значения Y_i решения $y(x)$ $(a\leqslant x\leqslant b)$ в узлах x_i $(i=1, 2, \ldots, n)$ будут определяться из системы

$$\lambda\sum_{j=1}^{n}A_jK_{ij}Y_j=f_i \quad (i=1, 2, \ldots, n). \tag{8}$$

Особенно просто применение метода конечных сумм для решения интегрального уравнения Вольтерра второго рода

$$y(x)-\lambda\int_a^x K(x, s)\,y(s)\,ds=f(x) \quad (a\leqslant x\leqslant b),$$

которое можно рассматривать как уравнение Фредгольма второго рода (см. § 1). Здесь $K_{ij}=0$ при $j>i$, и, следовательно,

соответствующая система (4) имеет вид

$$Y_i - \lambda \sum_{j=1}^{i} A_j K_{ij} Y_j = f_i \quad (i = 1, 2, \ldots, n). \tag{9}$$

Получилась линейная система с треугольной матрицей. Если

$$1 - \lambda A_i K_{ii} \neq 0 \quad (i = 1, 2, \ldots, n), \tag{10}$$

то из системы (9) последовательно находим:

$$Y_1 = f_1 (1 - \lambda A_1 K_{11})^{-1},$$
$$Y_2 = (f_2 + \lambda A_1 K_{21} Y_1)(1 - \lambda A_2 K_{22})^{-1},$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$Y_n = \left(f_n + \lambda \sum_{j=1}^{n-1} A_j K_{ij} Y_j\right)(1 - \lambda A_n K_{nn})^{-1}.$$

Условие (10) при данном λ заведомо выполнено, если коэффициенты A_j достаточно малы, чего всегда можно добиться.

Пример 1. Методом конечных сумм найти приближенное решение интегрального уравнения

$$y(x) + \int_0^1 x e^{xs} y(s)\, ds = e^x. \tag{11}$$

Выберем узлы $x_1 = 0$, $x_2 = \frac{1}{2}$, $x_3 = 1$. Значения ядра $K(x, s) = xe^{xs}$ и правой части $f(x) = e^x$ в соответствующих точках приведены в следующих таблицах:

Таблица значений K_{ij}

$s \backslash x$	0	1/2	1
0	0	0,5000	1
$\frac{1}{2}$	0	0,6420	1,6487
1	0	1,3592	2,7183

Таблица значений f_i

x_i	0	1/2	1
f_i	1	1,6487	2,7183

Используя квадратурную формулу Симпсона (см. [18])

$$\int_0^1 F(x)\,dx \approx \frac{1}{6}\left[F(0)+4F\left(\frac{1}{2}\right)+F(1)\right],$$

для определения приближенных значений Y_i $(i=1,\ 2,\ 3)$ решения $y(x)$ в узлах x_i получаем систему

$$Y_1=1,$$

$$Y_2+\frac{1}{6}(0{,}5000Y_1+2{,}5680Y_2+1{,}3592Y_3)=1{,}6487,$$

$$Y_3+\frac{1}{6}(Y_1+6{,}5948Y_2+2{,}7183Y_3)=2{,}7183,$$

или, после упрощений,

$$\left.\begin{aligned} Y_1&=1,\\ 1{,}4280Y_2+0{,}2265Y_3&=1{,}5654,\\ 1{,}0991Y_2+1{,}4531Y_3&=2{,}5516.\end{aligned}\right\} \qquad (12)$$

Решая систему (12), находим: $Y_1=1$; $Y_2=0{,}930$; $Y_3=1{,}053$. Приближенное решение можно выразить формулой

$$Y(x)=e^x-\frac{x}{6}(1+3{,}720e^{\frac{x}{2}}+1{,}053e^x).$$

Заметим, что точное решение уравнения (11) есть $y(x)=1$, как легко проверить непосредственно.

§ 6. Метод вырожденных ядер

Определение. Ядро $K(x,\ s)$ называется *вырожденным*, если оно может быть представлено в виде конечной суммы парных произведений:

$$K(x,\ s)=\sum_{i=1}^{n}\alpha_i(x)\,\beta_i(s), \qquad (1)$$

где функции $\alpha_i(x)$, так же как и функции $\beta_i(s)$ $(i=1,\ 2,\ \ldots,\ n)$, можно считать линейно независимыми.

Для таких ядер интегральное уравнение Фредгольма второго рода

$$y(x)=f(x)+\lambda\int_a^b K(x,\ s)\,y(s)\,ds \qquad (2)$$

решается весьма просто. Действительно, подставляя выражение (1) в уравнение (2), будем иметь

$$y(x)=f(x)+\lambda\sum_{i=1}^{n}c_i\alpha_i(x), \qquad (3)$$

где

$$c_i = \int_a^b \beta_i(s)\, y(s)\, ds \qquad (i = 1, 2, \ldots, n) \tag{4}$$

— некоторые постоянные коэффициенты. Если в выражение (4) подставить формулу (3), то для определения коэффициентов c_i $(i = 1, 2, \ldots, n)$ получим алгебраическую систему линейных уравнений

$$c_i = \int_a^b \beta_i(s) f(s)\, ds + \lambda \int_a^b \beta_i(s) \sum_{j=1}^n c_j \alpha_j(s)\, ds \qquad (i = 1, 2, \ldots, n)$$

или

$$c_i - \lambda \sum_{j=1}^n c_j \gamma_{ij} = f_i, \tag{5}$$

где

$$f_i = \int_a^b \beta_i(s) f(s)\, ds, \qquad \gamma_{ij} = \int_a^b \alpha_i(s)\, \beta_j(s)\, ds. \tag{6}$$

Систему (5) можно записать в виде

$$\sum_{j=1}^n (\delta_{ij} - \lambda\gamma_{ji})\, c_j = f_i \qquad (i = 1, 2, \ldots, n), \tag{5'}$$

где δ_{ij} — символ Кронекера.

Обозначим через $\Delta(\lambda)$ определитель системы (5'):

$$\Delta(\lambda) = \det(\delta_{ij} - \lambda\gamma_{ji}) \equiv \begin{vmatrix} 1-\lambda\gamma_{11} & -\lambda\gamma_{21} & \ldots & -\lambda\gamma_{n1} \\ -\lambda\gamma_{12} & 1-\lambda\gamma_{22} & \ldots & -\lambda\gamma_{n2} \\ \ldots & \ldots & \ldots & \ldots \\ -\lambda\gamma_{1n} & -\lambda\gamma_{2n} & \ldots & 1-\lambda\gamma_{nn} \end{vmatrix}$$

и через $\Delta_{ij}(\lambda)$ $(i, j = 1, 2, \ldots, n)$ — алгебраические дополнения соответствующих элементов $\delta_{ij} - \lambda\gamma_{ji}$ определителя $\Delta(\lambda)$.

Если $\Delta(\lambda) \neq 0$, то на основании правила Крамера [18] находим

$$c_i = \frac{\sum_{j=1}^n \Delta_{ji}(\lambda) f_j}{\Delta(\lambda)} \qquad (i = 1, 2, \ldots, n).$$

Следовательно, в силу (3) интегральное уравнение (2) имеет единственное решение

$$y(x) = f(x) + \lambda \sum_{i=1}^n \sum_{j=1}^n \frac{\Delta_{ji}(\lambda)}{\Delta(\lambda)} f_j \alpha_i(x). \tag{7}$$

Отсюда, подставляя вместо f_j соответствующее выражение (6) и заменяя сумму интегралов интегралом суммы, получим

$$y(x) = f(x) + \lambda \int_a^b \frac{\Delta(x, s, \lambda)}{\Delta(\lambda)} f(s)\, ds, \tag{8}$$

где

$$\Delta(x, s, \lambda) = \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i(x)\, \beta_j(s)\, \Delta_{ji}(\lambda).$$

Из формулы (8) вытекает, что функция

$$R(x, s, \lambda) = \frac{\Delta(x, s, \lambda)}{\Delta(\lambda)} = \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i(x)\, \beta_j(s)\, \frac{\Delta_{ji}(\lambda)}{\Delta(\lambda)} \tag{9}$$

есть резольвента интегрального уравнения (2).

Собственные значения ядра $K(x, s)$ определяются из уравнения

$$\Delta(\lambda) = 0. \tag{10}$$

Если λ_k $(k = 1, 2, \ldots, m;\ m \leqslant n)$ есть корень уравнения (10) (очевидно, $\lambda_k \neq 0$), то соответствующие собственные функции $\varphi_k(x)$ ядра $K(x, s)$, т. е. нетривиальные решения однородного уравнения

$$\tilde{y}(x) = \lambda_k \int_a^b K(x, s)\, \tilde{y}(s)\, ds,$$

будут иметь вид

$$\varphi_k(x) = \lambda_k \sum_{i=1}^{n} \tilde{c}_i^{(k)} \alpha_i(x),$$

где $\tilde{c}_i^{(k)}$ — ненулевые решения линейной однородной системы

$$\sum_{j=1}^{n} (\delta_{ij} - \lambda_k \gamma_{ji})\, \tilde{c}_j^{(k)} = 0 \qquad (i = 1, 2, \ldots, n).$$

Если $\lambda = \lambda_k$ есть собственное значение ядра $K(x, s)$, то неоднородное уравнение (2) или не имеет решений, или имеет бесконечно много решений.

Пример 1. Решить уравнение

$$y(x) = x^2 + \lambda \int_{-1}^{1} (x + s)\, y(s)\, ds. \tag{11}$$

Ядро $K(x, s) = x + s$ здесь, очевидно, вырожденное. Из уравнения (11) получаем

$$y(x) = x^2 + \lambda (c_1 x + c_2), \tag{12}$$

где

$$c_1=\int_{-1}^{1} y(s)\,ds, \quad c_2=\int_{-1}^{1} sy(s)\,ds. \tag{13}$$

Подставляя выражение (12) в формулы (13), будем иметь систему

$$c_1=\frac{2}{3}+2c_2\lambda, \qquad c_2=\frac{2}{3}\lambda c_1; \tag{14}$$

отсюда

$$c_1=\frac{\frac{2}{3}}{1-\frac{4}{3}\lambda^2}, \qquad c_2=\frac{\frac{4}{9}\lambda}{1-\frac{4}{3}\lambda^2}.$$

Следовательно, из формулы (12), если $\lambda^2\neq\frac{3}{4}$, получаем решение

$$y(x)=x^2+\frac{2\lambda x+\frac{4}{3}\lambda^2}{3-4\lambda^2}. \tag{15}$$

При $\lambda^2=\frac{3}{4}$ уравнение (11) решений не имеет.

Пример 2. Найти собственные значения, собственные функции и резольвенту ядра $K(x, s)=x+s$ в области $-1\leqslant x\leqslant 1$, $-1\leqslant s\leqslant 1$.

На основании однородного уравнения

$$y(x)=\lambda\int_{-1}^{1}(x+s)\,y(s)\,ds$$

имеем

$$y(x)=\lambda\,(\tilde{c}_1x+\tilde{c}_2), \tag{16}$$

где коэффициенты $\tilde{c}_1$ и $\tilde{c}_2$ определяются из системы (ср. (14))

$$\tilde{c}_1-2\lambda\tilde{c}_2=0, \qquad -\frac{2}{3}\lambda\tilde{c}_1+\tilde{c}_2=0. \tag{17}$$

Приравнивая нулю определитель системы

$$\Delta(\lambda)=\begin{vmatrix} 1 & -2\lambda \\ -\frac{2}{3}\lambda & 1 \end{vmatrix}, \tag{18}$$

получим уравнение

$$1-\frac{4}{3}\lambda^2=0,$$

из которого находим собственные значения: $\lambda_1=\frac{\sqrt{3}}{2}$ и $\lambda_2=-\frac{\sqrt{3}}{2}$. Так как ядро $K(x, s)$ симметрическое:

$$K(x, s)=K(s, x),$$

то собственные значения λ_1 и λ_2 действительны.

Из системы (17) имеем

$$\tilde{c}_1^{(k)} - 2\lambda_k \tilde{c}_2^{(k)} = 0 \qquad (k = 1, 2),$$

откуда

$$\frac{\tilde{c}_1^{(k)}}{2\lambda_k} = \frac{\tilde{c}_2^{(k)}}{1} = \tilde{c}^{(k)}.$$

Следовательно, на основании (16) собственные функции суть

$$\varphi_1(x) = u_1(\sqrt{3}\,x + 1), \qquad \varphi_2(x) = u_2(-\sqrt{3}\,x + 1),$$

где $u_k = \lambda_k \tilde{c}^{(k)} \neq 0 \qquad (k = 1, 2)$.

Собственные функции обычно нормируют, полагая

$$\int_a^b \varphi_k^2(x)\,dx = 1 \qquad (k = 1, 2).$$

В нашем случае имеем

$$u_k^2 \int_{-1}^{1} (\pm x\sqrt{3} + 1)^2\,dx = u_k^2 \cdot 4 = 1.$$

Отсюда можно принять $u_k = {}^1/_2$ $(k = 1, 2)$, и, следовательно, нормированные собственные функции суть

$$\tilde{\varphi}_1(x) = \frac{1}{2}(1 + x\sqrt{3}), \qquad \tilde{\varphi}_2(x) = \frac{1}{2}(1 - x\sqrt{3}).$$

Из определителя (18) получаем соответствующие алгебраические дополнения:

$$\Delta_{11}(\lambda) = 1, \qquad \Delta_{12}(\lambda) = \frac{2}{3}\lambda, \qquad \Delta_{21}(\lambda) = 2\lambda, \qquad \Delta_{22}(\lambda) = 1,$$

причем

$$\Delta(\lambda) = 1 - \frac{4}{3}\lambda^2.$$

Поэтому, учитывая, что

$$K(x, s) = \alpha_1(x)\beta_1(s) + \alpha_2(x)\beta_2(s),$$

где $\alpha_1(x) = x$, $\beta_1(s) = 1$, $\alpha_2(x) = 1$, $\beta_2(s) = s$, на основании формулы (9) находим резольвенту ядра:

$$R(x, s, \lambda) = \frac{1}{1 - \frac{4}{3}\lambda^2}\left(x + 2\lambda xs + \frac{2}{3}\lambda + s\right) = \frac{\frac{2}{3}\lambda + (x + s) + 2\lambda xs}{1 - \frac{4}{3}\lambda^2}.$$

Решение любого неоднородного уравнения

$$y(x) = f(x) + \lambda \int_{-1}^{1} (x + s)\,y(s)\,ds$$

при $\lambda^2 \neq \frac{3}{4}$ выражается формулой

$$y(x) = f(x) + \lambda \int_{-1}^{1} \frac{\frac{2}{3}\lambda + (x+s) + 2\lambda xs}{1 - \frac{4}{3}\lambda^2} f(s)\, ds.$$

Для приближенного решения общего интегрального уравнения

$$y(x) = f(x) + \lambda \int_a^b K(x, s)\, y(s)\, ds, \tag{19}$$

где функции $f(x)$ и $K(x, s)$ для простоты будем считать непрерывными, ядро $K(x, s)$ заменяют близким к нему вырожденным ядром

$$K^{(n)}(x, s) = \sum_{m=0}^{n} \alpha_m(x)\, \beta_m(s).$$

Укажем несколько способов такой замены. Если ядро $K(x, s)$ аналитическое по x на отрезке $[a, b]$, то в качестве вырожденного ядра $K^{(n)}(x, s)$ можно взять конечный отрезок ряда Тейлора:

$$K^{(n)}(x, s) = \sum_{m=0}^{n} \frac{(x - x_0)^m}{m!} K_{x^m}^{(m)}(x_0, s),$$

где x_0 — некоторая точка отрезка $[a, b]$. Аналогичный прием можно применить также, если $K(x, s)$ аналитично по s на отрезке $[a, b]$. Для построения вырожденного ядра можно также использовать конечный отрезок двойного ряда Тейлора (см. [20]):

$$K^{(n)}(x, s) = \sum_{p=0}^{n} \sum_{q=0}^{n} a_{pq} (x - x_0)^p (s - s_0)^q,$$

где

$$a_{pq} = \frac{1}{p!q!} \frac{\partial^{p+q}}{\partial x^p \partial x^q} K(x_0, s_0) \qquad (x_0, y_0 \in [a, b]).$$

Пусть $l = b - a$. Непрерывное ядро $K(x, s)$ допускает аппроксимацию тригонометрическим полиномом периода $2l$ (см. [20]). Например, можно положить

$$K^{(n)}(x, s) = \frac{1}{2} a_0(s) + \sum_{p=1}^{n} a_p(s) \cos \frac{p\pi x}{l}, \tag{20}$$

где $a_p(s)$ $(p = 0, 1, 2, \ldots)$ — коэффициенты Фурье:

$$a_p(s) = \frac{2}{l} \int_a^b K(x, s) \cos \frac{p\pi x}{l}\, dx. \tag{21}$$

Аналогичное разложение получается, если поменять ролями переменные x и s. Можно также использовать конечный отрезок двойного ряда Фурье. Полагая, например,

$$a_p(s) \approx \frac{1}{2} a_{p0} + \sum_{q=1}^{n} a_{pq} \cos \frac{q\pi s}{l}, \qquad (p = 0, 1, 2, \ldots, n)$$

из формул (20) и (21) имеем

$$K^{(n)}(x, s) = \frac{1}{4} a_{00} + \frac{1}{2} \sum_{p=1}^{n} a_{p0} \cos \frac{p\pi x}{l} + \frac{1}{2} \sum_{q=1}^{n} a_{0q} \cos \frac{q\pi s}{l} + \\ + \sum_{p=1}^{n} \sum_{q=1}^{n} a_{pq} \cos \frac{p\pi x}{l} \cos \frac{q\pi s}{l},$$

где

$$a_{pq} = \frac{4}{l^2} \int_a^b \int_a^b K(x, s) \cos \frac{p\pi x}{l} \cos \frac{q\pi s}{l} dx\, ds.$$

Наконец, полагая

$$h = \frac{b-a}{n}, \quad x_i = a + ih \qquad (i = 0, 1, 2, \ldots, n),$$

можно воспользоваться первым интерполяционным полиномом Ньютона (см. [18]) по x:

$$K^{(n)}(x, s) = K(x_0, s) + \sum_{m=1}^{n} \frac{(x-x_0)(x-x_1)\ldots(x-x_{m-1})}{m!h^m} \Delta^m K(x_0, s).$$

Аналогично можно взять первый интерполяционный полином Ньютона по s.

Иногда целесообразно использовать интерполяционный полином Ньютона для функции двух переменных [18]:

$$K^{(n)}(x, s) = K^{(n)}(x_0, x_0) + \\ + \sum_{p+q=1}^{p+q=n} \frac{\Delta_{x^p s^q}^{p+q} K(x_0, x_0)}{p!q!h^{p+q}} (x-x_0) \ldots (x-x_{p-1})(s-s_0)\ldots(s-s_{q-1}),$$

где $x_0 \in [a, b]$.

Употребляются и другие приемы интерполирования и аппроксимирования ядра $K(x, s)$.

Если $K^{(n)}(x, s)$ есть приближенное вырожденное ядро для точного ядра $K(x, s)$ и функция $f_n(x)$ также близка к $f(x)$, то

решение $z_n(x)$ интегрального уравнения

$$z_n(x) = f_n(x) + \lambda \int_a^b K^{(n)}(x, s)\, z_n(s)\, ds \qquad (n = 1, 2, \ldots) \tag{22}$$

можно рассматривать как приближение решения $y(x)$ уравнения (19).

Как показали Л. В. Канторович и В. И. Крылов [16], справедлива следующая оценка погрешности: пусть

$$\int_a^b |K(x, s) - K^{(n)}(x, s)|\, ds \leqslant \varepsilon, \qquad |f(x) - f_n(x)| \leqslant \delta$$

и резольвента $R_n(x, s, \lambda)$ для уравнения (22) такова, что

$$\int_a^b |R_n(x, s, \lambda)|\, ds \leqslant M_n(\lambda)$$

при $a \leqslant x \leqslant b$, причем выполнено неравенство

$$q = |\lambda|\, \varepsilon\, [1 + |\lambda|\, M_n(\lambda)] < 1;$$

тогда уравнение (19) имеет единственное решение $y(x)$ и

$$|y(x) - z_n(x)| \leqslant \frac{Nq\,[1 + |\lambda|\, M_n(\lambda)]}{1 - q} + \delta, \tag{23}$$

где $N \geqslant \max |f(x)|$ на $[a, b]$.

Из оценки (23) вытекает, что если $K^{(n)}(x, s) \rightrightarrows K(x, s)$ и $f_n(x) \rightrightarrows f(x)$ при $n \to \infty$, причем $M_n(\lambda) \leqslant M(\lambda) < +\infty$, то $z_n(x) \rightrightarrows y(x)$ на $[a, b]$.

Пример 3. Приближенно решить уравнение

$$y(x) - \int_0^{\frac{1}{2}} e^{-x^2 s^2}\, y(s)\, ds = 1. \tag{24}$$

Пользуясь известным разложением, ядро $K(x, s) = e^{-x^2 s^2}$ приближенно заменяем вырожденным ядром

$$K^{(2)}(x, s) = 1 - x^2 s^2 + \frac{x^4 s^4}{2}.$$

Отсюда вместо уравнения (24) получаем уравнение

$$z(x) = 1 + \int_0^{\frac{1}{2}} \left(1 - x^2 s^2 + \frac{x^4 s^4}{2}\right) z(s)\, dz. \tag{25}$$

Следовательно,

$$z(x) = 1 + c_1 + c_2 x^2 + c_3 x^4, \tag{26}$$

где

$$c_1=\int_0^{\frac{1}{2}} z(x)\,dx,\quad c_2=-\int_0^{\frac{1}{2}} x^2 z(x)\,dx,\quad c_3=\frac{1}{2}\int_0^{\frac{1}{2}} x^4 z(x)\,dx. \tag{27}$$

На основании формул (26) и (27) имеем систему

$$\left.\begin{aligned} c_1&=\frac{1}{2}+\frac{1}{2}c_1+\frac{1}{24}c_2+\frac{1}{160}c_3,\\ c_2&=-\frac{1}{24}-\frac{1}{24}c_1-\frac{1}{160}c_2-\frac{1}{896}c_3,\\ c_3&=\frac{1}{320}+\frac{1}{320}c_1+\frac{1}{1792}c_2+\frac{1}{9216}c_3. \end{aligned}\right\} \tag{28}$$

Решив систему (28), получим $c_1=0{,}9930$; $c_2=-0{,}0833$; $c_3=0{,}0007$. Следовательно,

$$y(x)\approx z(x)=1{,}9930-0{,}0833x^2+0{,}0007x^4 \quad \left(0\leqslant x\leqslant\frac{1}{2}\right). \tag{29}$$

Оценку погрешности приближенного решения (29) можно произвести по формуле (23), но это связано с громоздкими вычислениями.

§ 7. Метод коллокации

Рассмотрим интегральное уравнение

$$R[y]\equiv y(x)-\lambda\int_a^b K(x,\ s)\,y(s)\,ds-f(x)=0. \tag{1}$$

Будем искать приближенное решение уравнения (1) в виде функции определенного вида

$$Y_n=\Phi(x,\ c_1,\ c_2,\ \ldots,\ c_n) \tag{2}$$

со свободными параметрами $c_1, c_2, \ldots, c_n$ *(неопределенные коэффициенты)*. Подставляя выражение (2) в уравнение (1), получим *невязку*

$$R[Y_n]=Y_n(x)-\lambda\int_a^b K(x,\ s)\,Y_n(s)\,ds-f(x). \tag{3}$$

Если y является точным решением, то, очевидно, невязка $R[y]=0$. Поэтому стараются подобрать параметры $c_1, c_2, \ldots, c_n$ так, чтобы невязка $R[Y_n]$ была в определенном смысле возможно малой. Минимизировать невязку $R[Y_n]$ можно различными способами. Например, используют приемы, аналогичные тем, которые применяются для решения краевых задач дифференциальных уравнений (гл. IV, §§ 6—8). Здесь и в дальнейших параграфах мы рассмотрим некоторые методы минимизации невязки $R[Y_n]$, применяемые на практике.

Обычно для простоты выкладок берут функцию Y_n, линейно зависящую от параметров $c_1, c_2, \ldots, c_n$. Найдя параметры $c_1, c_2, \ldots, c_n$, получают приближенное решение (2).

Отметим одно обстоятельство. Если невязка $R[Y_n]$ получилась малой, то она близка к невязке $R[y]=0$, даваемой точным решением y. Но если два оператора $R[Y_n]$ и $R[y]$ имеют близкие значения, то отсюда, вообще говоря, не следует, что функции Y_n и y близки между собой в обычном смысле (например, в смысле равномерного приближения (гл. I, § 12)). Поэтому возникает математическая задача: по известной невязке $R[Y_n]$ оценить погрешность $|y-Y_n|$ приближенного решения Y_n. Эта трудная проблема связана с глубокими теоремами функционального анализа (см., например, [16], [22]), и мы ее оставим без рассмотрения, ограничившись лишь некоторыми указаниями частного характера.

Другой математический вопрос, который здесь также не затрагивается, — это вопрос сходимости Y_n к точному решению y при $n \to \infty$, т. е. выяснение условий, при которых имеет место предельное равенство

$$\lim_{n\to\infty} Y_n = y. \tag{4}$$

Если справедливо равенство (4), то данным методом решение y можно найти с любой степенью точности, взяв достаточно большое число параметров $c_1, c_2, \ldots, c_n$.

Перейдем теперь к изложению одного из конкретных методов построения приближенного решения Y_n.

Положим

$$Y_n(x) = \varphi_0(x) + \sum_{i=1}^{n} c_i\varphi_i(x), \tag{5}$$

где $\varphi_0(x), \varphi_1(x), \ldots, \varphi_n(x)$ — известные функции *(координатные функции)* и $c_1, c_2, \ldots, c_n$ — неопределенные коэффициенты, причем функции $\varphi_i(x)$ $(i=1, 2, \ldots, n)$ линейно независимы. Заметим, что, в частности, можно полагать $\varphi_0(x)\equiv 0$. Подставляя выражение (5) в левую часть уравнения (1), получим невязку

$$R[Y_n] = \varphi_0(x) + \sum_{i=1}^{n} c_i\varphi_i(x) - f(x) - \\ - \lambda\int_a^b K(x, s)\left[\varphi_0(s) + \sum_{i=1}^{n} c_i\varphi_i(s)\right] ds$$

или

$$R[Y_n(x)] = \psi_0(x, \lambda) + \sum_{i=1}^{n} c_i\psi_i(x, \lambda), \tag{6}$$

где

$$\left.\begin{aligned}\varphi_0(x, \lambda) &= \varphi_0(x) - f(x) - \lambda \int_a^b K(x, s)\,\varphi_0(s)\,ds,\\ \psi_i(x, \lambda) &= \varphi_i(x) - \lambda \int_a^b K(x, s)\,\varphi_i(s)\,ds \quad (i = 1, 2, \ldots, n).\end{aligned}\right\} \tag{7}$$

Согласно *методу коллокации* [23] (ср. гл. IV, § 9) требуем, чтобы невязка $R[Y_n(x)]$ обращалась в нуль в заданной системе точек x_j $(j = 1, 2, \ldots, n)$ из отрезка $[a, b]$ *(точки коллокации)*, т. е. полагаем, что

$$R[Y_n(x_j)] = 0 \quad (j = 1, 2, \ldots, n),$$

где $a \leqslant x_1 < x_2 < \ldots < x_{n-1} < x_n \leqslant b$.

Отсюда на основании формулы (6) для определения коэффициентов $c_1, c_2, \ldots, c_n$ получаем алгебраическую линейную систему уравнений

$$\sum_{i=1}^{n} c_i \psi_i(x_j, \lambda) = -\psi_0(x_j, \lambda) \quad (j = 1, 2, \ldots, n). \tag{8}$$

Если определитель системы (8) $D(\lambda) = \det[\psi_i(x_j, \lambda)] \neq 0$, то из системы (8) можно однозначно определить величины $c_1, c_2, \ldots, c_n$ и, следовательно, найти приближенное решение $Y_n(x)$ по формуле (5).

Приравнивая нулю определитель $D(\lambda)$, получим уравнение $D(\lambda) = 0$, из которого, вообще говоря, можно найти приближенные значения $\tilde{\lambda}_k$ $(k = 1, 2, \ldots, n)$ первых собственных чисел ядра $K(x, s)$.

Если положить $f(x) \equiv 0$, $\varphi_0(x) \equiv 0$, $\lambda = \tilde{\lambda}_k$, то вместо системы (8) будем иметь однородную систему

$$\sum_{i=1}^{n} \tilde{c}_i^{(k)} \psi_i(x_j, \tilde{\lambda}_k) = 0 \quad (j = 1, 2, \ldots, n). \tag{9}$$

Найдя ненулевые решения $\tilde{c}_i^{(k)}$ $(i = 1, 2, \ldots, n)$ системы (9), получим для ядра $K(x, s)$ приближенные собственные функции

$$\tilde{Y}_n^{(k)}(x) = \sum_{i=1}^{n} \tilde{c}_i^{(k)} \varphi_i(x),$$

отвечающие его собственному значению $\lambda_k \approx \tilde{\lambda}_k$.

Пример 1. Методом коллокации решить уравнение

$$y(x) - \int_0^1 \frac{s^2 y(s)}{x^2 + s^2}\,ds = x \operatorname{arctg} \frac{1}{x}. \tag{10}$$

Положим $Y(x) = c_1 + c_2 x$.

Подставляя это выражение в уравнение (10), получим невязку

$$R[Y(x)]=c_1 x \operatorname{arctg}\frac{1}{x}+c_2\left[x-\frac{1}{2}+\frac{x^2}{2}\ln\left(1+\frac{1}{x^2}\right)\right]-x\operatorname{arctg}\frac{1}{x}.$$

Выбирая точки коллокации $x_1=0$, $x_2=1$ и учитывая, что

$$\lim_{x\to 0} x\operatorname{arctg}\frac{1}{x}=0,\quad \lim_{x\to 0} x^2\ln\left(1+\frac{1}{x^2}\right)=0,$$

для определения коэффициентов c_1 и c_2 будем иметь систему

$$0\cdot c_1-\frac{1}{2}c_2=0,\qquad \frac{\pi}{4}c_1+\frac{1}{2}(1+\ln 2)c_2=\frac{\pi}{4}.$$

Отсюда получаем $c_2=0$, $c_1=1$. Таким образом,

$$Y=1. \tag{11}$$

Найденное приближенное решение (11), как легко проверить, является точным.

§ 8. Метод наименьших квадратов

Для уравнения

$$R[y]\equiv y(x)-\lambda\int_a^b K(x,s)\,y(s)\,ds-f(x)=0, \tag{1}$$

аналогично методу коллокации (§ 7), полагаем

$$Y_n(x)=\varphi_0(x)+\sum_{i=1}^{n} c_i\varphi_i(x), \tag{2}$$

где $\varphi_0(x)$, $\varphi_1(x)$, ..., $\varphi_n(x)$ — известные функции и c_1, c_2, ..., c_n — неопределенные коэффициенты, причем $\varphi_i(x)$ $(i=1, 2, \ldots, n)$ линейно независимы.

Подставляя (2) в левую часть уравнения (1), получим невязку

$$R[Y_n]=\psi_0(x,\lambda)+\sum_{i=1}^{n} c_i\psi_i(x,\lambda), \tag{3}$$

где $\psi_0(x,\lambda)$ и $\psi_i(x,\lambda)$ $(i=1, 2, \ldots, n)$ определяются формулами (7) § 7.

Согласно *методу наименьших квадратов* [16], [19], [24] (ср. гл. IV, § 7) коэффициенты c_i $(i=1, 2, \ldots, n)$ отыскиваются из условия минимума интеграла

$$I=\int_a^b \{R[Y_n]\}^2\,dx=\int_a^b\left[\psi_0(x,\lambda)+\sum_{i=1}^{n} c_i\psi_i(x,\lambda)\right]^2 dx. \tag{4}$$

Это требование приводит к алгебраической системе уравнений

$$\frac{\partial I}{\partial c_j}=0 \quad (j=1,\ 2,\ \ldots,\ n); \tag{5}$$

отсюда на основании (4), дифференцируя по параметрам $c_1, c_2, \ldots, c_n$ под знаком интеграла, будем иметь

$$\frac{1}{2}\frac{\partial I}{\partial c_j}=\int_a^b \psi_j(x,\ \lambda)\left[\psi_0(x,\ \lambda)+\sum_{i=1}^{n} c_i\psi_i(x,\ \lambda)\right]dx=0 \tag{6}$$

$$(j=1,\ 2,\ \ldots,\ n).$$

С помощью сокращенных обозначений

$$(\psi_i,\ \psi_j)=\int_a^b \psi_i(x,\ \lambda)\,\psi_j(x,\ \lambda)\,dx \tag{7}$$

систему (6) можно записать в виде *нормальной системы способа наименьших квадратов* (гл. I, § 6):

$$\left.\begin{array}{l} c_1(\psi_1,\ \psi_1)+c_2(\psi_1,\ \psi_2)+\ldots+c_n(\psi_1,\ \psi_n)=-(\psi_1,\ \psi_0),\\ c_1(\psi_2,\ \psi_1)+c_2(\psi_2,\ \psi_2)+\ldots+c_n(\psi_2,\ \psi_n)=-(\psi_2,\ \psi_0),\\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\ c_1(\psi_n,\ \psi_1)+c_2(\psi_n,\ \psi_2)+\ldots+c_n(\psi_n,\ \psi_n)=-(\psi_n,\ \psi_0). \end{array}\right\} \tag{8}$$

Заметим, что если $\varphi_0(x)\equiv 0$, то $\psi_0(x)=-f(x)$, и, следовательно, $-(\psi_i,\ \psi_0)=(\psi_i,\ f)$ $(i=1,\ 2,\ \ldots,\ n)$. Так как $(\psi_i,\ \psi_j)=(\psi_j,\ \psi_i)$, то матрица системы (8) симметрическая.

Вместо интегрального метода наименьших квадратов можно воспользоваться точечным способом наименьших квадратов (гл. I, § 3).

Метод наименьших квадратов применяется также для приближенного нахождения собственных значений и собственных функций ядра $K(x,\ s)$, аналогично тому как это делается для метода коллокации (§ 7). А именно, полагая $f(x)\equiv 0$ и $\varphi_0(x)\equiv 0$, откуда $\psi_0(x)\equiv 0$, определяем приближенные значения собственных чисел из алгебраического уравнения

$$\det[(\psi_i,\ \psi_j)]=0. \tag{9}$$

После этого приближенные собственные функции находятся из однородной системы (8), где вместо λ подставлено соответствующее приближенное значение.

Пример 1. Методом наименьших квадратов найти приближенное решение уравнения

$$y(x)=x^2+\int_{-1}^{1}\operatorname{sh}(x+s)\,y(s)\,ds. \tag{10}$$

Для первого приближения полагаем $Y=c_1+c_2x+x^2$. Отсюда

$$\varphi_1(x)=1, \quad \varphi_2(x)=x, \quad \varphi_0(x)=x^2.$$

Учитывая, что

$$\int_{-1}^{1} \operatorname{sh}(x+s)\,ds=a\operatorname{sh}x, \qquad \int_{-1}^{1} s\operatorname{sh}(x+s)\,ds=b\operatorname{sh}x, \qquad \int_{-1}^{1} s^2\operatorname{sh}(x+s)\,ds=c\operatorname{sh}x,$$

где $a=2\operatorname{sh}1=2{,}3504$; $b=2e^{-1}=0{,}7358$; $c=6\operatorname{sh}1-4\operatorname{ch}1=0{,}8788$, на основании формул (4) и (7) из § 7 имеем:

$$\psi_1=1-a\operatorname{sh}x, \quad \psi_2=x-b\operatorname{ch}x, \quad \psi_0=-c\operatorname{sh}x.$$

Далее находим:

$$(\psi_1, \psi_1)=2+a^2\left(\frac{\operatorname{sh}2}{2}-1\right)=6{,}4935,$$

$$(\psi_2, \psi_2)=\frac{2}{3}+b^2\left(\frac{\operatorname{sh}2}{2}+1\right)=2{,}1896,$$

$$(\psi_1, \psi_2)=-4\,(ae^{-1}+b\operatorname{sh}1)=-8e^{-1}\operatorname{sh}1=-3{,}4586,$$

$$(\psi_1, \psi_0)=ac\left(\frac{\operatorname{sh}2}{2}-1\right)=1{,}6800,$$

$$(\psi_2, \psi_0)=-2ce^{-1}=-0{,}6466.$$

Получаем систему для определения коэффициентов c_1 и c_2:

$$6{,}4935c_1-3{,}4586c_2=-1{,}6800; \qquad -3{,}4586c_1+2{,}1896c_2=0{,}6466.$$

Отсюда получаем: $c_1=-0{,}5423$; $c_2=-0{,}5613$. Таким образом,

$$Y=x^2-0{,}5613x-0{,}5423. \tag{11}$$

Так как для уравнения (10) ядро

$$K(x, s)=\operatorname{sh}(x+s)=\operatorname{sh}x\operatorname{ch}s+\operatorname{ch}x\operatorname{sh}s$$

вырожденное, то легко получить точное решение

$$y(x)=x^2+\alpha\operatorname{sh}x+\beta\operatorname{ch}x, \tag{12}$$

где

$$\alpha=\frac{6\operatorname{sh}1-4\operatorname{ch}1}{2-\left(\frac{\operatorname{sh}2}{2}\right)^2}=-0{,}6821, \qquad \beta=\alpha\left(\frac{\operatorname{sh}2}{2}-1\right)=-0{,}5548.$$

Из сравнения формул (11) и (12) заключаем, что приближенное решение Y близко к точному y, если $|x|$ — малая величина. На концах $x=\pm 1$ расхождение $|y-Y|$ довольно значительно.

§ 9. Метод моментов

Пусть

$$R[y]\equiv y(x)-\lambda\int_a^b K(x, s)\,y(s)\,ds-f(x)=0. \tag{1}$$

Аналогично предыдущему (§ 8), будем искать приближенное решение

уравнения (1) в виде конечной суммы

$$Y_n(x) = f(x) + \sum_{i=1}^{n} c_i \varphi_i(x) \qquad (i = 1, 2, \ldots n), \tag{2}$$

где $\varphi_i(x)$ $(i = 1, 2, \ldots, n)$ — некоторые известные линейно независимые функции *(координатные функции)* и $c_1, c_2, \ldots, c_n$ — неопределенные коэффициенты. Подставляя выражение (2) в левую часть уравнения (1), получим невязку

$$R[Y_n] = \sum_{j=1}^{n} c_j \left[\varphi_j(x) - \lambda \int_a^b K(x, s)\, \varphi_j(s)\, ds\right] - \lambda \int_a^b K(x, s) f(s)\, ds. \tag{3}$$

Согласно *методу моментов* [16], [19], [24] коэффициенты c_i $(i = 1, 2, \ldots, n)$ определяются из условия ортогональности невязки ко всем координатным функциям $\varphi_1(x), \varphi_2(x), \ldots, \varphi_n(x)$. Это дает систему уравнений

$$\int_a^b R[Y_n]\, \varphi_i(x)\, dx = 0 \qquad (i = 1, 2, \ldots, n)$$

или, в силу (3),

$$\sum_{j=1}^{n} c_j (\alpha_{ij} - \lambda \beta_{ij}) = \lambda \gamma_i \qquad (i = 1, 2, \ldots, n), \tag{4}$$

где

$$\alpha_{ij} = \int_a^b \varphi_i(x)\, \varphi_j(x)\, dx, \qquad \beta_{ij} = \int_a^b dx \int_a^b K(x, s)\, \varphi_i(x)\, \varphi_j(s)\, ds,$$

$$\gamma_i = \int_a^b dx \int_a^b K(x, s)\, \varphi_i(x) f(s)\, ds.$$

Если определитель системы (4) $D(\lambda) = \det(\alpha_{ij} - \lambda \beta_{ij})$ отличен от нуля, то из этой системы можно однозначно определить коэффициенты $c_1, c_2, \ldots, c_n$. Тогда формула (2) даст приближенное решение интегрального уравнения (1). Из уравнения $D(\lambda) = 0$ приближенно находятся собственные значения $\lambda_1, \lambda_2, \ldots, \lambda_n$ ядра $K(x, s)$. Найдя ненулевые решения однородной линейной системы

$$\sum_{j=1}^{n} \tilde{c}_j (\alpha_{ij} - \lambda_k \beta_{ij}) = 0 \qquad (i = 1, 2, \ldots, n),$$

легко построить (см. § 6) приближенные собственные функции $\tilde{y}^{(k)} x$, отвечающие данному собственному значению λ_k. Заметим, что метод моментов по идее совпадает с методом Галеркина (гл. IV, § 8).

Можно показать [16], [19], что метод моментов равносилен замене ядра $K(x, s)$ некоторым вырожденным ядром $K^{(n)}(x, s)$. Поэтому

для приближенного решения $Y_n(x)$ имеется оценка погрешности (см. [16], [19]), аналогичная приведенной выше (§ 6).

Пример 1. Найти первые два собственных значения интегрального уравнения

$$R[y] \equiv y(x) - \lambda \int_0^1 K(x, s)\, y(s)\, ds = 0,$$

где

$$K(x, s) = \begin{cases} s, & \text{если} \quad s \leqslant x, \\ x, & \text{если} \quad s > x. \end{cases} \tag{5}$$

На основании (5) имеем

$$R[y] = y(x) - \lambda \left\{ \int_0^x s y(s)\, ds + \int_x^1 x y(s)\, ds \right\}.$$

Положим $Y = c_1 x + c_2 x^2$. Тогда

$$R[Y] = c_1 x + c_2 x^2 - \lambda \left[\frac{c_1 x^3}{3} + \frac{c_2 x^4}{4} + x\left(\frac{c_1}{2} + \frac{c_2}{3}\right) - \left(\frac{c_1 x^3}{2} + \frac{c_2 x^4}{3}\right) \right] =$$
$$= c_1 \left[\left(1 - \frac{\lambda}{2}\right) x + \frac{\lambda}{6} x^3 \right] + c_2 \left[-\frac{\lambda}{3} x + x^2 + \frac{\lambda}{12} x^4 \right].$$

Ортогонализируя невязку $R[Y]$, будем иметь систему

$$\int_0^1 R[Y]\, x\, dx = 0, \quad \int_0^1 R[Y]\, x^2\, dx = 0$$

или

$$\left.\begin{aligned} c_1 \left[\left(1 - \frac{\lambda}{2}\right) \cdot \frac{1}{3} + \frac{\lambda}{6} \cdot \frac{1}{5} \right] + c_2 \left[-\frac{\lambda}{3} \cdot \frac{1}{3} + \frac{1}{4} + \frac{\lambda}{12} \cdot \frac{1}{6} \right] = 0, \\ c_1 \left[\left(1 - \frac{\lambda}{2}\right) \cdot \frac{1}{4} + \frac{\lambda}{6} \cdot \frac{1}{6} \right] + c_2 \left[-\frac{\lambda}{3} \cdot \frac{1}{4} + \frac{1}{5} + \frac{\lambda}{12} \cdot \frac{1}{7} \right] = 0. \end{aligned}\right\}$$

После упрощения получим систему

$$\left.\begin{aligned} c_1 (120 - 48\lambda) + c_2 (90 - 35\lambda) = 0, \\ c_1 (630 - 245\lambda) + c_2 (504 - 180\lambda) = 0. \end{aligned}\right\} \tag{6}$$

Приравнивая нулю определитель системы (6), получим уравнение для определения собственных значений:

$$D(\lambda) \equiv \begin{vmatrix} 120 - 48\lambda & 90 - 35\lambda \\ 630 - 245\lambda & 504 - 180\lambda \end{vmatrix} = 0.$$

Отсюда $65\lambda^2 - 1692\lambda + 3780 = 0$, или

$$\lambda^2 - 26{,}03\lambda + 58{,}15 = 0. \tag{7}$$

Из уравнения (7) будем иметь

$$\tilde{\lambda}_1 = 2{,}462; \quad \tilde{\lambda}_2 = 23{,}568.$$

Для сравнения укажем точные собственные значения:

$$\lambda_1 = \frac{\pi^2}{4} = 2{,}467 \quad \text{и} \quad \lambda_2 = \frac{9\pi^2}{4} = 22{,}206,$$

полученные из решения соответствующей краевой задачи:

$$y'' + \lambda y = 0, \quad y(0) = 0, \quad y'(1) = 0.$$

Таким образом, погрешность $\tilde{\lambda}_1$ равна примерно 0,2%, а $\tilde{\lambda}_2$—6%.

В заключение заметим, что рассмотренные методы минимизации невязки (§§ 7—9) применимы также к решению нелинейных интегральных уравнений. Имеются также другие методы решения интегральных уравнений, например метод Монте-Карло [25].

ЛИТЕРАТУРА К ГЛАВЕ VII

[1] Виарда Г., Интегральные уравнения, ГТТИ, 1933.
[2] Гурса Э., Курс математического анализа, т. III, ч. 2, ГТТИ, 1934, гл. XXX—XXXIII.
[3] Гюнтер Н. М., Основы математической физики, ч. 1, Интегральные уравнения, Кубуч, 1931.
[4] Курант Р., Гильберт Д., Методы математической физики, т. I, Гостехиздат, 1951, гл. III.
[5] Ловитт У. В., Линейные интегральные уравнения, ГТТИ, 1933.
[6] Михлин С. Г., Интегральные уравнения, изд. 2, Гостехиздат, 1949.
[7] Морс Ф. М., Фешбах Г., Методы математической физики, т. I, ИЛ, 1958, гл. 8.
[8] Мусхелишвили Н. И., Сингулярные интегральные уравнения, изд. 2, Физматгиз, 1962.
[9] Мюнц Г. М., Интегральные уравнения, т. I, ГТТИ, 1934.
[10] Петровский И. Г., Лекции по теории интегральных уравнений Гостехиздат, 1951.
[11] Привалов И. И., Интегральные уравнения, ОНТИ, 1935.
[12] Смирнов В. И., Курс высшей математики, т. 4, Гостехиздат, 1941, гл. II.
[13] Смирнов Н. С., Введение в теорию нелинейных интегральных уравнений, ОНТИ, 1936.
[14] Трикоми Ф., Интегральные уравнения, ИЛ, 1960.
[15] Уиттекер Е. Т., Ватсон Г. Н., Курс современного анализа, ГТТИ, 1933, ч. I, гл. 11.
[16] Канторович Л. В., Крылов В. И., Приближенные методы высшего анализа, изд. 4, Физматгиз, 1962.
[17] Маркушевич А. И., Краткий курс теории аналитических функций, изд. 2, «Наука», 1966, гл. IX.
[18] Демидович Б. П., Марон И. А., Основы вычислительной математики, изд. 3, «Наука», 1966, гл. VIII, XIV, XVI.
[19] Березин И. С., Жидков Н. П., Методы вычислений, т. I, т. II, Физматгиз, 1961.
[20] Фихтенгольц Г. М., Курс дифференциального и интегрального исчислений, Гостехиздат, 1948—1949, т. II, гл. II, т. III, гл. XIX.
[21] Уиттекер Э., Робинсон Г., Математическая обработка результатов наблюдений, ГТТИ, 1933, гл. X, XV.
[22] Михлин С. Г., Вариационные методы в математической физике, Гостехиздат, 1957.
[23] Коллатц Л., Численные методы решения дифференциальных уравнений, ИЛ, 1958, гл. V.
[24] Положий Г. Н. и др., Математический практикум, Физматгиз, 1960, гл. 7.
[25] Бут Э. Д., Численные методы, Физматгиз, 1959, гл. XII.

ПРИЛОЖЕНИЕ I

ОРТОГОНАЛЬНЫЕ ПОЛИНОМЫ ЧЕБЫШЕВА ДЛЯ $n+1$ РАВНООТСТОЯЩИХ ТОЧЕК ($\hat{P}_k(t) = \hat{P}_k(0) P_{k,n}(t)$)

$n=5$ (6 точек)

t	$\hat{P}_1(t)$	$\hat{P}_2(t)$	$\hat{P}_3(t)$	$\hat{P}_4(t)$	$\hat{P}_5(t)$
0	5	5	5	1	1
1	3	−1	−7	−3	−5
2	1	−4	−4	2	10
3	−1	−4	4	2	−10
4	−3	−1	7	−3	5
5	−5	5	−5	1	−1
$\sum \hat{P}^2(t)$	70	84	180	28	252

$n=6$ (7 точек)

t	$\hat{P}_1(t)$	$\hat{P}_2(t)$	$\hat{P}_3(t)$	$\hat{P}_4(t)$	$\hat{P}_5(t)$
0	3	5	1	3	1
1	2	0	−1	−7	−4
2	1	−3	−1	1	5
3	0	−4	0	6	0
4	−1	−3	1	1	−5
5	−2	0	1	−7	4
6	−3	5	−1	3	−1
$\sum \hat{P}^2(t)$	28	84	6	154	84

$n=7$ (8 точек)

t	$\hat{P}_1(t)$	$\hat{P}_2(t)$	$\hat{P}_3(t)$	$\hat{P}_4(t)$	$\hat{P}_5(t)$
0	7	7	7	7	7
1	5	1	−5	−13	−23
2	3	−3	−7	−3	17
3	1	−5	−3	9	15
4	−1	−5	3	9	−15
5	−3	−3	7	−3	−17
6	−5	1	5	−13	23
7	−7	7	−7	7	−7
$\sum \hat{P}^2(t)$	168	168	264	616	2184

$n=8$ (9 точек)

t	$\hat{P}_1(t)$	$\hat{P}_2(t)$	$\hat{P}_3(t)$	$\hat{P}_4(t)$	$\hat{P}_5(t)$
0	4	28	14	14	4
1	3	7	−7	−21	−11
2	2	−8	−13	−11	4
3	1	−17	−9	9	9
4	0	−20	0	18	0
5	−1	−17	9	9	−9
6	−2	−8	13	−11	−4
7	−3	7	7	−21	11
8	−4	28	−14	14	−4
$\sum \hat{P}^2(t)$	60	2772	990	2002	468

$n=9$ (10 точек)

t	$\hat{P}_1(t)$	$\hat{P}_2(t)$	$\hat{P}_3(t)$	$\hat{P}_4(t)$	$\hat{P}_5(t)$
0	9	6	42	18	6
1	7	2	−14	−22	−14
2	5	−1	−35	−17	1
3	3	−3	−31	3	11
4	1	−4	−12	18	6
5	−1	−4	12	18	−6
6	−3	−3	31	3	−11
7	−5	−1	35	−17	−1
8	−7	2	12	−22	14
9	−9	6	−42	18	−6
$\sum \hat{P}^2(t)$	330	132	8580	2860	780

$n=10$ (11 точек)

t	$\hat{P}_1(t)$	$\hat{P}_2(t)$	$\hat{P}_3(t)$	$\hat{P}_4(t)$	$\hat{P}_5(t)$
0	5	15	30	6	3
1	4	6	−6	−6	−6
2	3	−1	−22	−6	−1
3	2	−6	−23	−1	4
4	1	−9	−14	4	4
5	0	−10	0	6	0
6	−1	−9	14	4	−4
7	−2	−6	23	−1	−4
8	−3	−1	22	−6	1
9	−4	6	6	−6	6
10	−5	15	−30	6	−3
$\sum \hat{P}^2(t)$	110	858	4290	286	156

ПРИЛОЖЕНИЕ II

ПЕРВЫЕ 10 ПОЛИНОМОВ ЛЕЖАНДРА $P_n(x)$

n	$P_n(x)$
0	1
1	x
2	$\frac{1}{2}(3x^2-1)$
3	$\frac{1}{2}(5x^3-3x)$
4	$\frac{1}{8}(35x^4-30x^2+3)$
5	$\frac{1}{8}(63x^5-70x^3+15x)$
6	$\frac{1}{16}(231x^6-315x^4+105x^2-5)$
7	$\frac{1}{16}(429x^7-693x^5+315x^3-35x)$
8	$\frac{1}{128}(6435x^8-12\,012x^6+6930x^4-1260x^2+35)$
9	$\frac{1}{128}(12\,155x^9-25\,740x^7+18\,018x^5-4620x^3+315x)$
10	$\frac{1}{265}(46\,189x^{10}-109\,395x^8+90\,090x^6-30\,030x^4+3465x^2-63)$

ПРИЛОЖЕНИЕ III

ПЕРВЫЕ 12 ПОЛИНОМОВ ЧЕБЫШЕВА $T_n(x)$

$$T_1(x)=x$$
$$2T_2(x)=2x^2-1$$
$$4T_3(x)=4x^3-3x$$
$$8T_4(x)=8x^4-8x^2+1$$
$$16T_5(x)=16x^5-20x^3+5x$$
$$32T_6(x)=32x^6-48x^4+18x^2-1$$
$$64T_7(x)=64x^7-112x^5+56x^3-7x$$
$$128T_8(x)=128x^8-256x^6+160x^4-32x^2+1$$
$$256T_9(x)=256x^9-576x^7+432x^5-120x^3+9x$$
$$512T_{10}(x)=512x^{10}-1280x^8+1120x^6-400x^4+50x^2-1$$
$$1024T_{11}(x)=1024x^{11}-2816x^9+2816x^7-1232x^5+220x^3-11x$$
$$2048T_{12}(x)=2048x^{12}-6144x^{10}+6912x^8-3584x^6+840x^4-72x^2+1$$

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

Альтернанс чебышевский 78
Аппроксимация квадратичная 22
— — **интегральная** 40
— — **точечная** 22
Аппроксимирующий полином 12, 22

Базисные функции 232

Вариационная задача 308
— — , **метод Ритца** 321
Вычислительные шаблоны, построение 266

Гармоника 50
Гармоническая функция 255, 256
Гармонический анализ 51
Граница сетки 272

Дифференциальное уравнение второго порядка, приближенное решение 187, 189
— — — — **самосопряженное** 312
— —, **конечно-разностный метод решения** 151
— —, **метод Адамса** 156, 159, 160
— —, — — **для системы** 160
— —, — **Крылова последовательных сближений** 164, 167
— —, — **Милна** 169, 170, 174
— —, — — **для системы** 174
— — — **Ньютона — Канторовича** 201
— —, — **последовательных приближений** 134, 137, 138
— —, — — — **для системы** 138
— —, — **Рунге—Кутта** 151
— —, — — — **для системы** 152, 153
— —, — **Чаплыгина** 191
— —, — **численного интегрирования** 140
— —, — **Эйлера** 144, 145, 151
— —, **методы решения основанные на применении производных высших порядков** 181, 182, 183, 184, 186
— —, **модификации метода Эйлера** 147, 148, 151
— — **обыкновенное, методы решения** 127
— — — **первого порядка** 121
— —, **оценка погрешности приближенного решения** 202
— —, **решение с помощью рядов** 128, 130, 132
— —, **связь с уравнением Вольтерра** 336
— —, **усовершенствованный метод ломаных** 147
— —, — — **Эйлера — Коши** 148, 149
Дифференциальное уравнение *n*-**го порядка, сведение к системе** 122, 123

Задача вариационная 308
— **Дирихле** 254, 255, 320
— —, **единственность** 256
— —, **корректность** 256
— —, **метод моделирования** 270, 271
— —, — **Монте-Карло** 272, 274, 275
— —, — **Ритца** 329
— —, — **сеток** 261, 262, 269
— —, **решение с помощью сеточного электроинтегратора** 271
— **Коши** 121, 122, 124, 249, 252
— — **краевая, см.** *Краевая задача*
— — **общая** 250
— **Неймана** 254
— **о собственных значениях** 215
— **основная внешней баллистики** 125
— **смешанная** 251, 252
— — **корректно поставленная** 252
— — **некорректно поставленная** 253
— **трех тел** 126

Интеграл Дирихле 320
Интегральная кривая 123
— **поверхность** 244
Интегральное уравнение 332
— — **Вольтерра второго рода** 334, 335
— — — **первого рода** 334
— — **линейное** 332
— — **Фредгольма второго рода** 332, 335
— — — **однородное** 332
— — — **первого рода** 332
Интегральные уравнения, метод вырожденных ядер 345, 350
— —, — **конечных сумм** 341
Интерполяционный полином Лагранжа 13, 19
— — **Ньютона** 14
Интерполяция (интерполирование) 14, 23
— **квадратичная** 14
— **линейная** 14
— **тригонометрическая** 17
История блуждания 272
Итерированное ядро 339

Конечно-разностная схема неустойчивая 282
— — — **устойчивая** 281
Конечно-разностный оператор 279
Константа Липшица 121, 124
Корректность смешанной задачи 253
Коэффициенты Фурье 46
— — **полинома относительно системы ортогональных функций** 30

Коэффициенты Фурье тригонометрические 50
Краевая задача 254
— — внешняя, внутренняя 254
— — двухточечная простейшая 209
— — —, редукция к задаче Коши 217
— — для уравнения Лапласа 317
— — — — Пуассона 317
— — — — эллиптического типа 253
— — линейная 212, 213
— — — неоднородная, однородная 213
— — —, сведение к вариационной задаче 312
— — —, связь с интегральным уравнением Фредгольма 338
— —, метод Галеркина 238, 241
— —, — коллокации 232, 241
— —, — конечных разностей 219, 222, 223
— —, — наименьших квадратов 234, 235, 241
— —, — ортогональных проекций 242
— —, — подобластей 242
— —, — прогонки 224, 225, 226, 229
— —, — прямых 294, 297
— —, — сеток 261, 262, 269
— —, обобщенный метод Галеркина 242
— — общая для дифференциального уравнения 209
— — первая, вторая, третья 254
— — простейшая двухточечная 209
— — —, метод Ритца 322
— — смешанная 210, 254
— — —, условия устойчивости 282
— — Штурма—Лиувилля 324
— — — —, метод Ритца 325
Краевые условия 209, 247
— — двухточечные 213
— — линейные 213
— — точечно разделенные 217

Лемма Чаплыгина об интегральных неравенствах 192
Линеал 305
Линейная зависимость переменных 82
— — функций 28
— независимость функций 28
Линейный дифференциальный оператор 306
Ломаная Эйлера 144

Метод Адамса для дифференциальных уравнений 156, 159, 160
— —, оценка погрешности 206
— вариационный для краевых задач 309
— выбранных точек 93
— выравнивания 85, 88
— вырожденных ядер для интегральных уравнений 345
— Галеркина для краевых задач 238, 241
— — обобщенный 242
— графический ломаных Эйлера 128
— коллокации для краевых задач 232, 241
— — — уравнения Фредгольма 354
— конечных разностей (метод сеток) 261
— — — для дифференциальных уравнений обыкновенных 151
— — — — краевых задач 219, 222, 223
— — сумм для интегральных уравнений 341, 342, 343
— Крылова последовательных сближений 164, 167
— ломаных усовершенствований 147
Метод Милна решения дифференциальных уравнений 169, 170, 174
— моделирования для задачи Дирихле 271
— моментов для уравнения Фредгольма 359
— Монте-Карло для задачи Дирихле 272, 274, 275
— наименьших квадратов, см. *Способ наименьших квадратов*
— Ньютона — Канторовича для решения задачи Коши 201
— общий определения параметров эмпирической формулы 115
— ортогональных проекций для краевых задач 242
— подобластей для краевых задач 242
— последовательных приближений 134, 137, 138
— — — для интегральных уравнений 339, 340
— — — для краевых задач 224, 225, 226, 229
— — — для уравнения теплопроводности 285, 286
— прямых для уравнений эллиптического типа 293, 297
— — — уравнения Пуассона 297
— Ритца для вариационной задачи 321
— — — задачи Дирихле 328
— — — — Штурма — Лиувилля 325
— — — простейшей краевой задачи 322
— Рунга — Кутта для решения дифференциальных уравнений 151, 152, 153
— — —, оценка погрешности 206
— сеток для задачи Дирихле 261, 262, 269
— — для уравнений гиперболического типа 290
— средних для отыскания параметров эмпирической формулы 94
— Чаплыгина решения дифференциальных уравнений 191
— численного интегрирования дифференциальных уравнений 140
— Эйлера решения дифференциальных уравнений 145, 147, 151
— — с итерационной обработкой ординат 149
— Эйлера—Коши усовершенствованный 148, 149
Многочлен — см. *Полином*
Множество расчетных точек связное 263
— функций линейное 305

Начальная таблица 184
Начальные данные Коши 249
— условия 247
Начальный отрезок 157
Невязка 92, 238, 240
Неравенство Бесселя 47
Норма вектора 175
— функции 27, 44

Область задания оператора 305
— — функционала 304
— определения функционала 304
— распространения решения задачи Коши 251
Оператор 305
— дифференцирования 306
— Лапласа 247, 306
— линейный 306
— — дифференциальный 306
— положительный 307
— симметричный 307

Определитель Вандермонда 13
Ортогональная система функций 27
Ортогональность с весом 63
— системы интегрируемых функций 44
Ортогональные полиномы Чебышева, таблица 362
— функции 27
Ортонормированная система функций 27
Особые точки резольвенты 339
Остаточный член интерполяционного полинома Лагранжа 15
Отклонение абсолютное обобщенного полинома от функции 73
— среднее квадратичное функций 71
Оценка погрешности метода Адамса 206
— — — Рунге—Кутта 206
— — приближенного решения дифференциального уравнения 202
Ошибка аппроксимации 279

Погрешность метода 204
— текущая 204
Полином аппроксимирующий 12
— интерполяционный 13
— — Лагранжа 13, 15
— — Ньютона 14
— наилучшего равномерного приближения 74
— наименее отклоняющийся от нуля 75
— — — от функции 74
— обобщенный 12
—, разложение по системе ортогональных функций 29
— тригонометрический 12, 17
— Фурье 33
— — тригонометрический 50
— Чебышева, корни 68
— — ортогональный 34, 35
Полиномы Лежандра 56, 364
— —, производящая функция 60
— —, рекуррентная формула 61
— —, усиленное свойство ортогональности 59
— Чебышева 65, 75
— —, максимальное значение модуля 69
— — ортогональные 34, 35
— — — нормированные 35
— —, рекуррентные формулы 66
— —, свойство ортогональности 66
— —, экстремальное свойство 68
Порядок точности приближенного метода 204
Приближение интегральное 12
— равномерное функции обобщенным полиномом 73
— точечное 12
— функций, постановка задачи 12
Приближенное решение вариационной задачи 321
— — дифференциальных уравнений, см. *Дифференциальное уравнение, метод*
— — краевых задач, см. *Краевая задача, метод*
Принцип максимума для гармонических функций 255
— Рунге 205
Процесс усреднения Либмана 265, 266

Равенство Парсеваля 48
Расчетные точки 263
Резольвента уравнения Фредгольма 339
Решение системы обыкновенных дифференциальных уравнений 123
— уравнения с частными производными 244
Ряд Фурье тригонометрический 51

Самосопряженный вид линейного дифференциального уравнения 312
Сеточная область 261
Символ Кронекера 44
Система интегрируемых функций ортогональная 44
— — — ортонормированная 44
— нормальная линейных уравнений 96
— обыкновенных дифференциальных уравнений 122
— ортогональных полиномов Чебышева нормированная 35
— — функций неполная 48
— — — полная 48
— уравнений метода прямых 295
— условных уравнений 116
— функций ортогональная 27
— — — с весом 63
— — ортонормированная 27
Скалярное произведение функций 307
Слой фиктивный 290
Случайное блуждание равномерное 272
Собственная функция 215
Собственное значение 215
Собственные значения (числа) задачи Штурма — Лиувилля 324
— — интегрального уравнения 333
— решения задачи Штурма — Лиувилля 325
— функции задачи Штурма — Лиувилля 324
— — интегрального уравнения 333
Способ наименьших квадратов для краевых задач 234, 235, 241
— — — для случая обобщенного полинома 25
— — — для уравнения Фредгольма 356
— — — для эмпирической формулы 96, 97
— — — интегральный 40
— — — точечный 22, 23, 24
Схема вычислений неустойчивая 206
— — устойчивая 206, 240
— неявная 286
— явная 279

Теорема Вейерштрасса аппроксимационная 74
— единственности тригонометрического полинома 18
— Неймана 256
— о корнях полинома Лежандра 60
— — — — Чебышева 68
— о полиноме Чебышева 75
— о свойстве ортогональности полиномов Чебышева 66
— об ортогональности полиномов Лежандра 57
— об экстремальных значениях полиномов Чебышева 68
— существования и единственности полинома, обладающего минимальным отклонением 75
— — — — решения задачи Коши 124
— Чебышева об альтернансе 77
Теоремы основные вариационного метода решения краевой задачи 309, 310
— Фредгольма 333
Точки коллокации 233, 355

Траектория частицы 272
Тригонометрический полином Фурье 50
— ряд Фурье 51

Узел внутренний 262
— граничный 263
— интерполирования 13
— первого, второго ряда 263
Узлы сетки соседние 262
Уклонения (невязки) 92, 93
Уравнение Абеля обобщенное 334
— волновое 246
— Вольтерра первого, второго рода 334, 335
— —, метод конечных сумм 341, 343
— —, — последовательных приближений 340
— —, связь с дифференциальным уравнением 336
— гиперболического типа, метод сеток 290
— интегральное, см. *Интегральное уравнение*
— колебаний струны 246
— — —, метод сеток 290, 292
— Лапласа 247, 306
— — в конечных разностях 257
— параболического типа, решение методом сеток 278
— Пуассона 247, 297, 306
— —, метод прямых 297
— с частными производными вполне линейное 244
— — — — гиперболического типа 245
— — — —, классификация 245
— — — — параболического типа 245
— — — — смешанного типа 245
— — — — эллиптического типа 245
— теплопроводности 245
— —, решение методом прогонки 285, 289
— —, устойчивость конечно-разностной схемы 281
— Фредгольма 332, 338, 345
— —, метод вырожденных ядер 345
— —, — коллокации 354
— —, — конечных сумм 342, 343
— —, — моментов 359
— —, — наименьших квадратов 356
— —, — последовательных приближений 338
— характеристик 246
— частот задачи Штурма — Лиувилля 325
— эллиптического типа, метод прямых 293, 297
Уравнения условные 116
Условие Липшица 121, 124
— полноты системы 48
— эллиптичности 294
Условия краевые 247
— начальные 121, 122, 124, 247

Формула Адамса экстраполяционная 157
— горизонтальной строки 164
— ломаной строки первая, вторая 164
— Милна контрольная 172
— — первая, вторая 171
— наклонной строки 164
— Родрига 56
— типа Рунге—Кутта 152
— эмпирическая 80
Функции базисные 232
— координатные 328, 354
— линейно зависимые 28
— — независимые 28
— ортогональные 27
Функционал 304
— линейный 305
Функция влияния 337
— гармоническая 255, 256
— Грина 337
— допустимая 304, 305
— периодическая, интерполирование тригонометрическим полиномом 17

Характеристики уравнения с частными производными 246
Характеристическое уравнение задачи Штурма — Лиувилля 325
— число 215
Ход метода прогонки обратный 225, 287
— прямой 225, 287

Чебышевский альтернанс 78

Эмпирическая формула, квадратичная (параболическая) зависимость 89, 91
— —, критерий прямолинейности ряда точек 83
— —, линейная зависимость 82
— —, метод выбранных точек 93
— —, — выравнивания 85, 88
— —, — наименьших квадратов 96, 97
— —, — средних 94
— —, общий метод определения параметров 115
— —, определение параметров 92, 93, 94, 96
— —, построение 80
— —, проверка пригодности 103
— — с двумя параметрами 101
— — с тремя параметрами 108, 109
— —, уточнение 112
Эмпирические формулы 102
— —, простейшие необходимые условия 102, 103

Ядро интегрального уравнения вырожденное 345
— — — симметричное 333
— итерированное 339