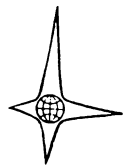
Р. БЕЛЛМАН
Р. КАЛАБА

КВАЗИЛИНЕАРИЗАЦИЯ И НЕЛИНЕЙНЫЕ КРАЕВЫЕ ЗАДАЧИ



ИЗДАТЕЛЬСТВО
„МИР“

QUASILINEARIZATION AND NONLINEAR BOUNDARY-VALUE PROBLEMS

Richard E. Bellman
and
Robert E. Kalaba
The RAND Corporation

AMERICAN ELSEVIER PUBLISHING COMPANY, INC.
New York 1965

Р. Беллман, Р. Калаба

КВАЗИЛИНЕАРИЗАЦИЯ И НЕЛИНЕЙНЫЕ КРАЕВЫЕ ЗАДАЧИ

ПЕРЕВОД С АНГЛИЙСКОГО
И. А. Вателя и Ф. И. Ерешко

ПОД РЕДАКЦИЕЙ
Ф. Л. Черноусько

ИЗДАТЕЛЬСТВО „МИР“
Москва 1968

УДК 517.91+518.61+519.3

Книга посвящена одному из методов прикладной математики — методу квазилинеаризации.

На простом примере уравнения Риккати авторы показывают характерные свойства вычислительного процесса, основанного на применении идей квазилинеаризации. Эти свойства в дальнейшем прослеживаются при исследовании более сложных случаев. В заключение рассматриваются нелинейные эллиптические и параболические уравнения и различные конкретные задачи физики, техники, механики и биологии.

Книга представляет интерес не только для математиков, но и для научных работников других специальностей и инженеров, имеющих дело с методами вычислительной математики. Она доступна студентам и аспирантам соответствующих специальностей.

Редакция литературы по математическим наукам

Инд. 2-2-3

Предисловие редактора перевода

Предлагаемая читателям книга известных американских математиков Р. Беллмана и Р. Калабы посвящена разработанному ими методу квазилинеаризации и его разнообразным приложениям. Метод квазилинеаризации, с помощью которого решение нелинейной задачи сводится к решению последовательности линейных задач, представляет собой по существу дальнейшую разработку известного метода Ньютона и его обобщенного варианта, предложенного Л. В. Канторовичем. Авторы приводят ряд результатов по сходимости метода; при этом широко используется аппарат дифференциальных неравенств, основные идеи которого принадлежат С. А. Чаплыгину.

Круг проблем, затронутых в этой небольшой книге, чрезвычайно широк. Авторы рассматривают различные математические задачи, к которым применяется аппарат квазилинеаризации: двухточечные и многоточечные краевые задачи для линейных и нелинейных обыкновенных дифференциальных уравнений, краевые задачи для эллиптических и параболических уравнений в частных производных, вариационные задачи, дифференциально-разностные и функционально-дифференциальные уравнения и др. Большое внимание в книге уделяется не только вопросам сходимости методов, но и реализации их на вычислительных машинах. Авторы приводят много числовых примеров, а в конце книги даны программы на языке ФОРТРАН для решения некоторых рассмотренных в тексте задач [1]).

Широко представлены и приложения метода квазилинеаризации к некоторым техническим, физическим и биологическим задачам (вопросы управления и идентификации систем, определение орбит, задачи переноса излучения, вопросы кардиологии).

---

[1]) Ссылки на программы имеются в соответствующих главах и параграфах книги. В тексте самих программ есть комментарии, поясняющие смысл используемых в них обозначений и ход работы программ. В русском издании программы воспроизведены с английского оригинала полностью, английский текст внутри программ сохранен. Мы надеемся, что это не помешает читателю пользоваться программами.

Сочетание строгости математических построений с арсеналом прикладных приемов и рецептов их машинной реализации является важным достоинством книги.

Следует отметить, что для многих из рассмотренных в книге задач в настоящее время развиты и другие численные методы решения, и было бы интересно сравнить с ними метод квазилинеаризации на примерах достаточно сложных многомерных задач.

Как и другие книги Р. Беллмана, уже известные советским читателям в русских переводах, эта книга написана живо и увлекательно, содержит немало свежих идей и интересных параллелей. Местами, однако, изложение авторов становится слишком беглым, а часть доказательств проводится лишь для простейших случаев. Некоторые прикладные задачи разбираются весьма схематично и могут рассматриваться лишь в качестве иллюстративных моделей.

В целом книга Р. Беллмана и Р. Калабы представляет собой актуальное и оригинальное исследование по одной из важнейших и труднейших проблем вычислительной математики — численному решению нелинейных краевых задач. Книга, несомненно, заинтересует специалистов в области дифференциальных уравнений, вычислительной математики, теории управления и их разнообразных приложений.

Ф. Л. Черноусько

Введение

Современные электронные вычислительные машины могут обеспечить быстрое и точное численное решение систем, состоящих из тысячи нелинейных обыкновенных дифференциальных уравнений, если задан полный набор начальных условий. Поэтому, как только физическая, экономическая, техническая или биологическая проблема сведена к задаче Коши для обыкновенных дифференциальных уравнений, можно считать, что ее полное решение близко к завершению. Следовательно, перед математиком стоит великая цель отыскания путей построения необходимых теорий и постановок задач, приближений и редукций.

Иногда, как, например, в теории движения планет, основными являются обыкновенные дифференциальные уравнения. В этом случае, однако, некоторые из постановок фундаментальных проблем не содержат полного набора начальных условий. Фактически основная задача определения орбит состоит в нахождении всей совокупности начальных условий по заданным угловым измерениям, сделанным в различные моменты времени. В других областях, в частности в математической физике и математической биологии, основными уравнениями являются дифференциальные уравнения в частных производных, интегро-дифференциальные уравнения, дифференциально-разностные уравнения и уравнения еще более сложного типа. Для того чтобы полностью использовать те мощные вычислительные возможности, которыми мы сейчас располагаем, необходимо проделать большой предварительный анализ постановок задач, а также аналитических методов, которые имеет смысл применять. Если раньше стремились так упростить задачи, чтобы получить линейные функциональные уравнения, в настоящее время целью является преобразование вычислительных проблем к задачам Коши для обыкновенных дифференциальных уравнений (линейных или нелинейных). Теории динамического программирования и инвариантного погружения в целом ряде областей успешно справляются с этой целью — успех достигается путем введения новых переменных состояния и использования полугрупповых свойств в пространстве, времени и структуре. Квазилинеаризация

достигает этой цели с помощью различных искусных комбинаций методов линейного приближения и использования возможностей цифровых машин. Последовательные приближения строятся таким образом, чтобы добиться быстрой сходимости, а по возможности и монотонности процесса.

Свое начало квазилинеаризация берет в теории динамического программирования. Существуют также связи с классической геометрией и анализом, но совершенно иного характера, и они не описываются никакой простой схемой. Одной из основ квазилинеаризации является геометрическая теория двойственности, связанная с именами Минковского, Фенхеля, Куна, Таккера и многих других; имеется также тесная связь с двойственностью, введенной в вариационное исчисление Фридрихсом и развитой с тех пор очень многими авторами. В другом направлении просматривается связь с богатым наследием теории дифференциальных неравенств, основанной Чаплыгиным и продолжаемой очень активной русской школой. Тесно примыкает к этому современная теория положительных и монотонных операторов, созданная Коллатцем и значительно развитая Редхеффером и др. Она в свою очередь связана с теорией обобщенной выпуклости, которой занимаются Беккенбах, Бонсалл, Пиксото и ряд других аналитиков. Что касается многочисленных приложений, то, по-видимому, квазилинеаризация есть просто применение метода Ньютона — Рафсона — Канторовича в функциональном пространстве.

Не останавливаясь подробно ни на одной из этих очень интересных областей, каждая из которых заслуживает монографий, мы упоминаем в тексте лишь связи с ними и даем интересующемуся читателю ряд ссылок для дальнейшего ознакомления.

Легко указать цели теории квазилинеаризации. Во-первых, мы хотим иметь единообразный подход к исследованию существования и единственности решений обыкновенных дифференциальных уравнений и уравнений в частных производных, подчиненных начальным и граничным условиям. Кроме того, мы хотим получить теоремы о представлении этих решений через решения линейных уравнений. Наконец, мы хотим дать единый подход к численным решениям как дескриптивных, так и вариационных задач, который обладает рядом свойств монотонности, а также быстротой сходимости.

Везде в этой книге, когда мы используем термин «численное решение», мы имеем в виду алгоритмы, которые должны реализоваться на цифровых вычислительных машинах.

Тот факт, что один и тот же идейный и аналитический подход служит всем поставленным целям, доставляет эстетическое

удовольствие, явля ся желательным с педагогической и эффективным с научной точки зрения. Эта книга, задуманная как введение в квазилинеаризацию, адресована и тем, кто интересуется только анализом, и тем, кто главным образом занимается приложениями.

Перейдем теперь к краткому описанию содержания семи глав. Чтобы познакомить читателя с аналитическими методами в их простейшем виде, мы начинаем гл. I с применения квазилинеаризации к исследованию уравнения Риккати

$$u' + u^2 + a_1(t)\,u + a_2(t) = 0. \tag{1}$$

Среди обыкновенных дифференциальных уравнений это уравнение одно из наиболее важных как само по себе, так и благодаря его связи с линейным дифференциальным уравнением второго порядка

$$w'' + a_1(t)\,w' + a_2(t)\,w = 0. \tag{2}$$

Мы покажем, как квазилинеаризация позволяет получить аналитическое представление решения u через квадратуры и операцию максимизации. Отсюда можно будет получить оценки сверху и снизу для u и w.

В гл. II мы перейдем к обсуждению двухточечных краевых задач для уравнений вида

$$u'' = g(u',\ u,\ t) \tag{3}$$

с граничными условиями

$$h_1(u'(0),\ u(0)) = 0, \quad h_2(u'(1),\ u(1)) = 0. \tag{4}$$

Мы изучим квадратичную сходимость последовательности приближений, порождаемой квазилинеаризацией, и приведем некоторые численные примеры.

Глава III посвящена исследованию условий, при которых последовательность приближений сходится монотонно. Это приводит к изучению линейного дифференциального неравенства вида

$$u'' + p(t)\,u' + q(t)\,u \geqslant 0. \tag{5}$$

Здесь мы столкнемся с функциями Грина. Чтобы рассмотреть свойство знакопостоянства этих функций, мы используем ряд интересных методов, таких, как факторизация оператора, вариационные методы, разностная аппроксимация уравнений в частных производных. Эти вопросы тесно связаны с задачами на собственные значения.

В гл. IV рассматривается несколько сходных задач о монотонности для дифференциальных уравнений более высокого порядка и для линейных систем. Изложенные здесь аналитические

результаты весьма отрывочны, и даже их не так легко получить. Мы приведем ряд примеров вычислительной эффективности методов.

Тщательное исследование эффективности численных методов показывает, что общепринятое использование последовательных приближений может привести к затруднениям с машинной памятью, когда порядок дифференциального уравнения или системы высок. В таком случае мы вынуждены разрабатывать методы использования последовательных приближений, не возлагающие слишком больших надежд на возможности машинной памяти, и вообще исследовать заманчивую область хранения и передачи информации. Мы предлагаем очень простой метод, позволяющий обойти трудности, связанные с хранением информации, и затем этот метод применяем к численному решению дифференциально-разностных уравнений вида

$$u'(t) = g(u(t),\ u(t-t_1),\ t) \tag{6}$$

и функционально-дифференциальных уравнений вида

$$u'(t) = g(u(t),\ u(h(t)),\ t). \tag{7}$$

Уравнения этого типа играют важную роль в современной математической биологии.

Дальнейшие усилия по преодолению трудностей, связанных с хранением информации, приводят к методу дифференциального приближения. Этот метод был разработан в сотрудничестве с И. М. Ричардсоном из Исследовательских лабораторий Хьюза в Малибу с целью получить аналитические приближения для различных классов сложных функциональных уравнений. Общая постановка задачи для случая дифференциальных уравнений имеет следующий вид.

Для заданного векторного уравнения

$$\frac{dx}{dt} = g(x,\ t), \quad x(0) = c, \tag{8}$$

мы хотим найти более удобное для анализа уравнение вида

$$\frac{dy}{dt} = h(y,\ t,\ a), \quad y(0) = b, \tag{9}$$

причем так, чтобы $y \cong x$.

Здесь векторы a и b должны быть выбраны так, чтобы соответствующим образом определенная норма $\|x - y\|$, например $\int_0^T (x-y,\ x-y)\,dt$, была мала. Эта вариационная задача поддается решению методом квазилинеаризации. В качестве при-

менения данного подхода мы показываем, как с помощью конечной системы обыкновенных дифференциальных уравнений получить точное решение уравнения восстановления, имеющего вид

$$u(t) = 1 + \int_0^t e^{-(t-s)^2} u(s)\, ds. \tag{10}$$

В гл. V мы приводим результаты приложений квазилинеаризации к решению нелинейного эллиптического уравнения в частных производных

$$u_{xx} + u_{yy} = e^u \tag{11}$$

и нелинейного параболического уравнения в частных производных

$$u_t = u_{xx} + g(u), \tag{12}$$

подчиненных разнообразным начальным и граничным условиям.

Хотя из-за больших затрат времени на вычисления исследовано только несколько случаев, имеются ясные указания на то, что использование квазилинеаризации для численного решения уравнений в частных производных должно дать выигрыш как во времени расчетов, так и в их точности. Однако еще многое предстоит сделать, прежде чем можно будет сформулировать такие же четкие утверждения, как в случае обыкновенных дифференциальных уравнений.

Глава VI содержит ряд примеров, иллюстрирующих некоторые пути, на которых может найти применение квазилинеаризация для решения дескриптивных и вариационных задач, возникающих в планировании и управлении, при определении орбит, при выявлении периодичностей и при идентификации систем. Последняя тема охватывает обратные задачи переноса излучения и кардиологии.

Заключительная глава посвящена динамическому программированию. Она начинается с пояснения пути, по которому приближение в пространстве политик приводит к квазилинеаризации. Затем указывается, как можно использовать динамическое программирование для того, чтобы избежать двухточечных краевых задач, возникающих при решении вариационных задач с квадратичными функционалами. Глава заканчивается обсуждением вопроса о том, как можно использовать локальную квазилинеаризацию и глобальное динамическое программирование для расширения областей применимости обеих теорий.

Как всегда, мы опирались на советы и мнения друзей и коллег. В этой связи мы особенно хотим выразить благодарность Т. А. Брауну, О. Гроссу, М. Янкозу, Р. Шридару. Наш замысел

оказался бы пустой фантазией, не будь помощи со стороны И. Бюлля, Б. Коткина и особенно Х. Кагивады, заключавшейся в составлении программ для вычислительных машин и в некоторых мелочах, которые и делают различие между успехом и удачей. Наконец, мы выражаем нашу признательность Джанетт Блад, чье терпение, трудолюбие, смекалка и чувство юмора позволили нам завершить нелегкий труд по написанию этой книги.

Все сделанное нами является прологом к тому, что еще предстоит сделать в будущем. Мы надеемся, что другие найдут материал для исследования в этих областях, такой же интересный и благодарный, какой был у нас, и мы также надеемся, что наши методы, только опробованные нами, будут культивироваться более широко.

Глава I

Уравнение Риккати

1. Введение

В этой главе мы хотим применить методы аппроксимации к изучению одного из основных уравнений математического анализа, а именно нелинейного дифференциального уравнения первого порядка

$$u' + u^2 + p(t)\,u + q\,(t) = 0, \tag{1.1}$$

называемого уравнением Риккати. Это уравнение играет фундаментальную роль при анализе поведения решений линейного дифференциального уравнения второго порядка

$$w'' + p(t)\,w' + q(t)\,w = 0 \tag{1.2}$$

и тем самым занимает важное положение в квантовой механике в связи с изучением уравнения Шредингера. Кроме того, вместе со своими многомерными аналогами оно занимает центральное место в динамическом программировании и инвариантном погружении.

Рассмотрение уравнения (1.1) даст нам возможность проиллюстрировать применение методов квазилинеаризации без каких-либо арифметических и алгебраических усложнений. Особенно просто можно установить основные свойства — монотонность и квадратичную сходимость, и, таким образом, можно без труда получить образец для большинства излагаемых далее задач.

2. Метод Ньютона — Рафсона

Для пояснения подхода, который мы будем неоднократно использовать в этой книге, рассмотрим сначала задачу нахождения корня скалярного уравнения

$$f(x) = 0 \tag{2.1}$$

методом последовательных приближений. Будем предполагать, что $f(x)$ — монотонно убывающая строго выпуклая функция для всех x из области определения, т. е. $f''(x) > 0$ (рис. 1.1). Следовательно, корень r — простой, $f'(r) \neq 0$.

Пусть x_0 — начальное приближение для корня r, причем $x_0 < r$ $(f(x_0) > 0)$, и пусть $f(x)$ аппроксимируется линейной функ-

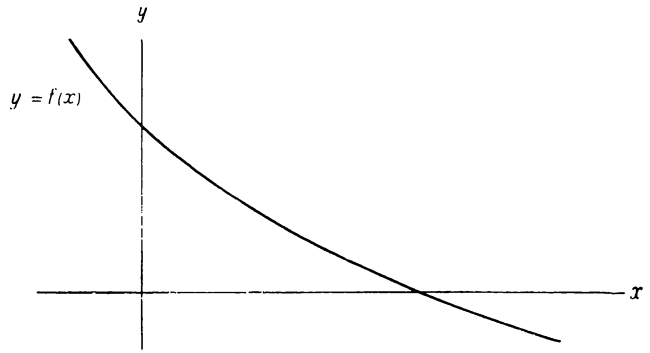


Рис. 1.1.

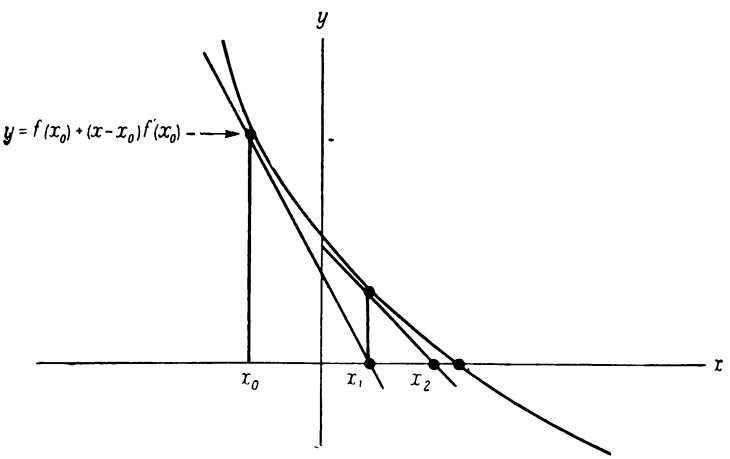


Рис. 1.2.

цией от x, задаваемой значениями самой функции $f(x)$ и ее производной при $x = x_0$ (рис. 1.2):

$$f(x) \cong f(x_0) + (x - x_0) f'(x_0). \tag{2.2}$$

Тогда следующее приближение для r получается из решения *линейного* по x уравнения

$$f(x_0) + (x - x_0) f'(x_0) = 0. \tag{2.3}$$

Отсюда имеем второе приближение

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}. \tag{2.4}$$

Затем процесс повторяется для x_1, давая новое значение x_2 и т. д., как показано на рис. 1.2. Общее рекуррентное соотношение имеет вид

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}. \tag{2.5}$$

Из рис. 1.2 видно, что

$$x_0 < x_1 < x_2 < \ldots < r, \tag{2.6}$$

если вначале $x_0 < r$. Аналитически это следует из неравенств $f(x_n) > 0$, $f'(x_n) < 0$. Такая монотонность с вычислительной точки зрения является важным свойством. Разумеется, мы предпочли бы иметь поочередное приближение сверху и снизу от r, т. е. чтобы было $x_0 < x_2 < x_4 < \ldots < r < \ldots < x_3 < x_1$, но этого обычно очень трудно добиться.

Второе свойство, даже более важное для вычислительного процесса, возможно, не так очевидно. Это так называемая *квадратичная сходимость*. Утверждается, что при некоторых разумных предположениях относительно $f'''(x)$

$$|x_{n+1} - r| \leqslant k |x_n - r|^2, \tag{2.7}$$

где k не зависит от n.

Чтобы убедиться в справедливости (2.7), запишем

$$x_{n+1} - r = x_n - \frac{f(x_n)}{f'(x_n)} - r = x_n - \frac{f(x_n)}{f'(x_n)} - \left(r - \frac{f(r)}{f'(r)}\right) = \varphi(x_n) - \varphi(r), \tag{2.8}$$

где $\varphi(x) = x - f(x)/f'(x)$. С помощью формулы Тейлора, используя первые три ее члена, включая остаточный, получаем

$$x_{n+1} - r = (x_n - r)\varphi'(r) + \frac{(x_n - r)^2}{2}\varphi''(\theta), \tag{2.9}$$

где $x_0 \leqslant x_n \leqslant \theta \leqslant r$. Поскольку

$$\varphi'(x) = \frac{f(x) f''(x)}{f'(x)^2}, \tag{2.10}$$

мы видим, что $\varphi'(r) = 0$.

Следовательно,

$$|x_{n+1} - r| \leqslant k |x_n - r|^2, \tag{2.11}$$

где $k = \max\limits_{x_0 \leqslant \theta \leqslant r} |\varphi''(\theta)|/2$ зависит от $f'''(x)$.

Кроме того,

$$x_{n+1}-x_n=\varphi(x_n)-\varphi(x_{n-1})= \\ =(x_n-x_{n-1})\varphi'(x_{n-1})+\frac{(x_n-x_{n-1})^2}{2}\varphi''(\theta), \quad (2.12)$$

где $x_{n-1}\leqslant 0\leqslant x_n$, откуда получаем

$$(x_{n+1}-x_n)=(x_n-x_{n-1})\frac{f(x_{n-1})f''(x_{n-1})}{f'(x_{n-1})^2}+\frac{(x_n-x_{n-1})^2\varphi''(\theta)}{2}. \quad (2.13)$$

Согласно (2.5), заменим в выражении (2.13) величину $f(x_{n-1})/f'(x_{n-1})$ на $x_{n-1}-x_n$. Тогда получим

$$(x_{n+1}-x_n)=(x_n-x_{n-1})^2\left[-\frac{f''(x_{n-1})}{f'(x_{n-1})}+\frac{\varphi''(\theta)}{2}\right]. \quad (2.14)$$

Таким образом,

$$|x_{n+1}-x_n|\leqslant k_1|x_n-x_{n-1}|^2, \quad (2.15)$$

где

$$k_1=\max_{x_0\leqslant\theta\leqslant r}\left[\frac{|f''(\theta)|}{|f'(\theta)|}+\frac{|\varphi''(\theta)|}{2}\right]. \quad (2.16)$$

Свойство, выражаемое соотношением (2.15), называется *квадратичной сходимостью.*

Очевидно, сходимость значительно ускоряется, когда x_n приближается к r. Асимптотически получается, что каждый следующий шаг удваивает число правильных знаков в данном приближении. Это свойство оказывается особенно полезным в задачах, содержащих громоздкие действия, не только потому, что время расчета, как правило, прямо пропорционально числу итераций, но также и потому, что ошибка округления может сильно возрастать с возрастанием числа итераций. Интересно проследить аналогию между рассматриваемым методом вычисления и суммированием асимптотического расходящегося ряда. Существует оптимальная частичная сумма, для которой ошибка минимальна.

3. Многомерный вариант

Метод без труда обобщается на случай многих переменных. Рассмотрим задачу нахождения решения системы уравнений

$$f_i(x_1, x_2, \ldots, x_N)=0, \quad i=1, 2, \ldots, N, \quad (3.1)$$

или в векторном обозначении

$$f(x)=0. \quad (3.2)$$

Пусть x_0 — начальное приближение, тогда

$$f(x)\cong f(x_0)+J(x_0)(x-x_0), \quad (3.3)$$

где $J(x_0)$ — матрица Якоби,

$$J(x_0)=\left(\frac{\partial f_i}{\partial x_j}\right)_{x=x_0}. \tag{3.4}$$

Новое приближение определяется по формуле

$$x_1=x_0-J(x_0)^{-1}f(x_0), \tag{3.5}$$

и мы получаем общее рекуррентное соотношение

$$x_n=x_{n-1}-J(x_{n-1})^{-1}f(x_{n-1}). \tag{3.6}$$

При соответствующих предположениях относительно функций $f(x)$ и их частных производных нетрудно теперь получить результаты, сходные с результатами предыдущего раздела.

Аналогичные результаты имеют место и для функциональных уравнений, однако их изложение требует более сложных формулировок и доказательств.

4. Квадратные корни

В качестве довольно интересного приложения метода, изложенного в разд. 2, рассмотрим задачу нахождения положительного решения уравнения $x^2=y$, $y>0$, т. е. задачу извлечения квадратного корня из y. Положим $f(x)=x^2-y$; тогда рекуррентное соотношение примет вид

$$x_{n+1}=x_n-\frac{(x_n^2-y)}{2x_n}=\frac{x_n}{2}+\frac{y}{2x_n}. \tag{4.1}$$

Это хорошо известный алгоритм, описанный Героном Александрийским.

Доказательство монотонности и квадратичной сходимости никоим образом не исчерпывает всех вопросов, возникающих при использовании этого алгоритма. Пусть нам нужно извлечь квадратные корни из ряда положительных чисел $y_1, y_2, \ldots, y_N$, лежащих в фиксированном интервале $[1, a]$. Предположим, что мы должны заранее выбрать начальное приближение x_0, не зная, какие конкретные значения примут эти числа. Как наилучшим образом выбрать фиксированное начальное приближение?

Один из возможных путей точной формулировки этого вопроса состоит в том, что мы потребуем минимизации функции

$$\max_{1\leqslant y\leqslant a}|x_N-\sqrt{y}| \tag{4.2}$$

для заданного значения x_0 при условии, что итерации (4.1) производятся ровно N раз. Пусть $x_0(N)$ есть значение, минимизирующее выражение (4.2). Можно показать, что

$$\lim_{N\to\infty}x_0(N)=a^{1/4}. \tag{4.3}$$

Этот результат получили Сильвестр и Хаммер. Аналогичные результаты можно получить для алгоритма, связанного с задачей извлечения корня k-й степени, $k > 1$. Здесь мы имеем

$$x_{n+1} = \frac{k-1}{k} x_n + \frac{y}{k x_n^{k-1}}. \tag{4.4}$$

Соответствующий предел для $x_0(N)$ можно легко угадать, но он еще строго не получен. Аналогичные задачи для более общих алгебраических, трансцендентных и функциональных уравнений, по-видимому, не только не исследовались, но даже и не ставились.

5. Квазилинеаризация

Рассмотрим теперь уравнение

$$x + bx^n = a, \quad a, \ b > 0, \quad n > 1, \tag{5.1}$$

и покажем, как методы квазилинеаризации позволяют получить выражение для единственного положительного решения с помощью операции взятия максимума.

Сначала заметим, что

$$x^n = \max_{y \geqslant 0} [nxy^{n-1} - (n-1)y^n] \tag{5.2}$$

для $x \geqslant 0$ и $n > 1$, причем максимум достигается при $y = x$. Впоследствии мы рассмотрим способ получения подобных соотношений, а сейчас примем (5.2) как легко проверяемое тождество.

Тогда (5.1) можно переписать в виде

$$x + \max_{y \geqslant 0} [bnxy^{n-1} - b(n-1)y^n] = a, \tag{5.3}$$

так как $b > 0$. Из (5.3) следует, что

$$x + bnxy^{n-1} - b(n-1)y^n \leqslant a \tag{5.4}$$

для всех $y \geqslant 0$. Следовательно,

$$x \leqslant \frac{a + b(n-1)y^n}{1 + bny^{n-1}} \tag{5.5}$$

для всех $y \geqslant 0$, так как $1 + bny^{n-1} > 0$. Поскольку равенство в (5.5) достигается при $y = x$, можно написать

$$x = \min_{y \geqslant 0} \left[\frac{a + b(n-1)y^n}{1 + bny^{n-1}} \right]. \tag{5.6}$$

Это есть точное аналитическое представление искомого решения. Выражения такого типа очень полезны, так как позволяют путем соответствующего выбора y быстро найти верхнюю границу для решения.

6. Уравнение Риккати

Применим теперь метод квазилинеаризации к уравнению Риккати. Нижеследующий анализ имеет не только иллюстративное значение, но и полезен сам по себе, потому что, как упоминалось выше, уравнения Риккати вместе со своими аналогами в многомерном и в функциональном пространствах играют фундаментальную роль в современной теории управления и в математической физике, особенно в связи с динамическим программированием и инвариантным погружением.

Уравнение Риккати — это нелинейное дифференциальное уравнение первого порядка вида

$$v' + v^2 + p(t)\,v + q(t) = 0. \tag{6.1}$$

Несмотря на довольно простой вид, это уравнение неразрешимо в квадратурах для произвольных коэффициентов $p(t)$ и $q(t)$.

На первый взгляд связь между уравнением Риккати и линейным дифференциальным уравнением второго порядка кажется чисто формальной. В уравнении

$$u'' + p(t)\,u' + q(t)\,u = 0 \tag{6.2}$$

делается замена переменных

$$u = e^{\int v\,dt}, \quad v = \frac{u'}{u} \tag{6.3}$$

и получается уравнение (6.1) относительно v.

Как читатель увидит впоследствии при анализе вариационных задач с помощью метода динамического программирования, связь между двумя указанными уравнениями гораздо глубже, значительней и многообразней.

7. Линейное уравнение первого порядка

Прежде чем начинать изучение уравнения (6.1) с помощью квазилинеаризации и метода последовательных приближений, напомним решение линейного уравнения первого порядка

$$\frac{dw}{dt} - f(t)\,w = g(t), \quad w(0) = c. \tag{7.1}$$

Умножая (7.1) на интегрирующий множитель $e^{-\int_0^t f(s)\,ds}$, получаем уравнение

$$\frac{d}{dt}\left(we^{-\int_0^t f(s)\,ds}\right) = g(t)\,e^{-\int_0^t f(s)\,ds}, \tag{7.2}$$

которое легко интегрируется.

Следовательно, решение уравнения (7.1) определяется выражением

$$w = ce^{\int_0^t f(s)\,ds} + e^{\int_0^t f(s)\,ds} \int_0^t g(s) e^{-\int_0^s f(r)\,dr}\,ds, \tag{7.3}$$

которое нам понадобится в дальнейшем.

Вместо этой довольно громоздкой формулы напишем выражение

$$w = T(f, g), \tag{7.4}$$

определяющее $T(f, g)$ как линейную неоднородную операцию над функцией g. Так как экспоненциальная функция положительна, то из условия $g_1(t) \geqslant g_2(t)$ для $t \geqslant 0$ следует, что $T(f, g_1) \geqslant T(f, g_2)$. Это очень важное и полезное свойство является основой для получения последующих результатов данной главы.

8. Решение уравнения Риккати с помощью операции взятия максимума

Покажем теперь, что уравнение Риккати можно решить с помощью операции взятия максимума. Рассмотрим параболу

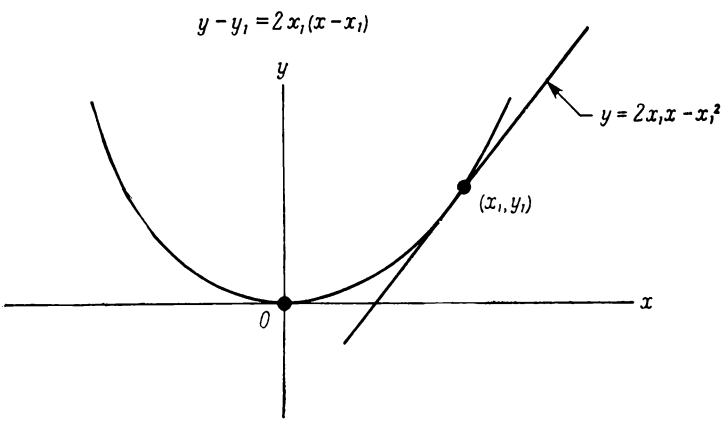


Рис. 1.3.

$y = x^2$ и касательную к ней в точке $(x_1, y_1) = (x_1, x_1^2)$ (рис. 1.3):

$$y - y_1 = 2x_1(x - x_1), \tag{8.1}$$

или

$$y = 2x_1x - x_1^2. \tag{8.2}$$

Так как кривая $y=x^2$ в силу выпуклости лежит выше касательной, то справедливо неравенство

$$x^2 \geqslant 2x_1x - x_1^2 \tag{8.3}$$

для всех x_1, причем равенство достигается только при $x=x_1$. Следовательно,

$$x^2 = \max_{x_1} [2x_1x - x_1^2], \tag{8.4}$$

что можно непосредственно установить многими способами. Эта простая геометрическая идея лежит в основе результатов (5.2) и (4.4).

Возвращаясь к уравнению (6.1)

$$v' = -v^2 - p(t)v - q(t), \tag{8.5}$$

заменим величину v^2 эквивалентным ей выражением

$$\max_u [2uv - u^2],$$

так что (8.5) примет вид

$$v' = -\max_u [2uv - u^2] - p(t)v - q(t) = \min_u [u^2 - 2uv - p(t)v - q(t)]. \tag{8.6}$$

Если бы не операция взятия минимума, это уравнение было бы линейным. Тем не менее оно обладает некоторыми свойствами, присущими линейным уравнениям. Именно по этим причинам мы используем термин «квазилинеаризация».

Рассмотрим соответствующее уравнению (8.6) линейное дифференциальное уравнение

$$w' = u^2 - 2uw - p(t)w - q(t), \tag{8.7}$$

где $u(t)$ — теперь уже *фиксированная* функция от t. Пусть v и w имеют одинаковые начальные значения $v(0)=w(0)=c$.

Предположим, что решение $v(t)$ уравнения Риккати существует на отрезке $[0, t_0]$. Оговорка относительно существования необходима, так как решение уравнения Риккати не обязательно существует при всех t. Это видно на примере уравнения $u'=1+u^2$, $u(0)=0$, решение которого $u=\operatorname{tg} t$ определено при $0 \leqslant t < \pi/2$.

Чтобы установить неравенство $w \geqslant v$, мы рассуждаем, как и выше. Заметим, что (8.6) эквивалентно неравенству

$$v' \leqslant u^2 - 2uv - p(t)v - q(t) \tag{8.8}$$

для *всех* $u(t)$ или равенству

$$v' = u^2 - 2uv - p(t)v - q(t) - r(t), \tag{8.9}$$

где $r(t) \geqslant 0$ для $t \geqslant 0$. Функция $r(t)$, разумеется, зависит от u и v, но это в данный момент несущественно.

Следовательно, используя обозначение разд. 7 и свойство положительности оператора $T(f, g)$, получаем

$$v = T(-2u - p(t),\ u^2 - q(t) - r(t)) \leqslant$$
$$\leqslant T(-2u - p(t),\ u^2 - q(t)) = w. \quad (8.10)$$

Так как это неравенство справедливо для всех $u(t)$, причем равенство достигается при $u(t) = v(t)$, мы получаем искомый результат [1])

$$v(t) = \min_u T(-2u - p(t),\ u^2 - q(t)), \quad (8.11)$$

который мы сформулируем в виде следующей теоремы.

Теорема 1. *Для $0 \leqslant t \leqslant t_0$, где t_0 зависит от интервала существования $v(t)$, имеет место представление*

$$v(t) = \min_u \left\{ e^{-\int_0^t (2u(s) + p(s))\, ds} \times \right.$$
$$\left. \times \left[c + \int_0^t e^{\int_0^s (2u(r) + p(r))\, dr} [u^2(s) - q(s)]\, ds \right] \right\}. \quad (8.12)$$

Итак, используя квазилинеаризацию, мы получили аналитическое выражение для решения уравнения Риккати.

9. Оценки сверху и снизу

С помощью представления (8.12) можно путем соответствующего выбора $u(t)$ легко получить оценки сверху для $v(t)$. Отсюда в свою очередь можно получить оценки сверху для решения линейного дифференциального уравнения второго порядка

$$u'' + p(t)u' + q(t)u = 0, \quad (9.1)$$

связь которого с уравнением Риккати отмечалась в разд. 6.

Если $q(t) < 0$, можно получить оценки снизу, сделав в (6.1) замену $v = 1/w$ и произведя описанные выше операции над уравнением

$$w' - 1 - p(t)w - q(t)w^2 = 0. \quad (9.2)$$

[1]) В соотношениях (8.11) и (8.12) минимум берется по функциям $u(t)$, заданным на отрезке $[0, t_0]$. — *Прим. ред.*

Эти методы использовались, особенно Калогеро, для изучения асимптотического поведения решения как функции от параметра и от времени t.

Далее, чтобы получить более точные оценки, можно использовать общие преобразования вида

$$v = \frac{a_1(t) + a_2(t)\, w}{b_1(t) + b_2(t)\, w}. \tag{9.3}$$

Функция w будет удовлетворять уравнению Риккати, если таким свойством обладает функция v.

10. Последовательные приближения с помощью квазилинеаризации

Теперь мы в состоянии использовать формулу теоремы 1 из разд. 8 для получения последовательных приближений к решению $v(t)$. Нам известно, что минимум достигается на самом решении $v(t)$. Следовательно, если v_0 — разумное начальное приближение, то v_1, получаемое как решение уравнения

$$v_1' = v_0^2 - 2v_0 v_1 - p(t)\, v_1 - q(t), \quad v_1(0) = c, \tag{10.1}$$

будет еще лучшим приближением, точно так же, как это имело место в процессе нахождения корня уравнения $f(x) = 0$ с помощью метода Ньютона — Рафсона. Повторяя этот процесс, мы получаем рекуррентное соотношение

$$v_{n+1}' = v_n^2 - 2v_n v_{n+1} - p(t)\, v_{n+1} - q(t), \quad v_{n+1}(0) = c. \tag{10.2}$$

В точности то же самое рекуррентное соотношение мы получили бы, если бы применили метод приближений Ньютона — Рафсона — Канторовича к нелинейному дифференциальному уравнению

$$v' = -v^2 - p(t)\, v - q(t), \quad v(0) = c. \tag{10.3}$$

Вообще, если бы дифференциальное уравнение имело вид $v' = g(v, t)$, мы бы использовали аппроксимационную схему

$$v_{n+1}' = g(v_n, t) + (v_{n+1} - v_n)\, g_v(v_n, t), \quad v_{n+1}(0) = c. \tag{10.4}$$

Мы хотим, однако, подчеркнуть, что эти два подхода неэквивалентны, несмотря на то что в некоторых случаях они приводят к одинаковым результатам. Как указывалось во введении (а также будет подробно обсуждаться в гл. VII), квазилинеаризация своим появлением во многом обязана теории динамического программирования, в которой концепция приближения в пространстве политик играет ведущую роль. В частности, с этой точки зрения мы можем ожидать монотонной сходимости.

С другой стороны, геометрическая трактовка метода приближений Ньютона — Рафсона позволяет ожидать квадратичной сходимости. Как будет показано ниже, оба эти свойства имеют место, и в результате мы получаем метод, весьма мощный как с вычислительной, так и с аналитической точек зрения.

Наконец, мы можем рассматривать квазилинеаризацию как систематическое применение принципа двойственности.

11. Монотонность

Докажем теперь монотонность последовательности приближений. Рассмотрим соотношения

$$v'_{n+1} = v_n^2 - 2v_n v_{n+1} - p(t)\,v_{n+1} - q(t) \geqslant$$
$$\geqslant v_{n+1}^2 - 2v_{n+1}v_{n+1} - p(t)\,v_{n+1} - q(t), \quad v_{n+1}(0) = c \quad (11.1)$$

(неравенство имеет место, так как минимум достигается при $v_n = v_{n+1}$). Мы умышленно не упростили последнее выражение, чтобы подчеркнуть его связь с уравнением

$$v'_{n+2} = v_{n+1}^2 - 2v_{n+2}v_{n+1} - p(t)\,v_{n+2} - q(t), \quad v_{n+2}(0) = c. \quad (11.2)$$

Сравнивая это уравнение с неравенством (11.1) и используя свойство линейного оператора T, рассмотренное в разд. 7, мы видим, что [1])

$$v_{n+1} \geqslant v_{n+2}. \quad (11.3)$$

Отсюда по индукции имеем

$$v_1 \geqslant v_2 \geqslant \ldots \geqslant v_n \geqslant v_{n+1} \geqslant \ldots. \quad (11.4)$$

Если функция v_0 выбрана произвольно, то неравенство $v_0 \geqslant v_1$ может и не выполняться.

Чтобы показать, что $v_n(t) \geqslant v(t)$ для всех $n \geqslant 1$, напишем

$$v' = -v^2 - p(t)\,v - q(t) = \min_w [w^2 - 2wv - p(t)\,v - q(t)] \leqslant$$
$$\leqslant v_n^2 - 2v_n v - p(t)\,v - q(t). \quad (11.5)$$

Отсюда опять, используя свойство линейного оператора (см. разд. 7), получаем, что $v \leqslant v_n$ для всех $n \geqslant 1$.

Следовательно, для любого отрезка $[0, t_0]$, на котором $v(t)$ существует, мы можем утверждать, что последовательность $\{v_n\}$ сходится.

---

[1]) Уравнение (11.1) можно записать в виде (7.1), полагая $w = v_{n+1}$, $f = -2v_{n+1} - p$, $g_1 = (v_n - v_{n+1})^2 + v_{n+1}^2 - q$. Сравнивая его с уравнением (11.2), которое также приводится к виду (7.1) при $w = v_{n+2}$, $f = -2v_{n+1} - p$, $g_2 = v_{n+1}^2 - q \leqslant g_1$, приходим к неравенству (11.3). — *Прим. ред.*

Из соотношения

$$v_{n+1}(t) - c = \int_0^t [v_n^2 - 2v_n v_{n+1} - p(t_1) v_{n+1} - q(t_1)]\, dt_1 \qquad (11.6)$$

заключаем на основе ограниченной сходимости, что

$$\bar{v}(t) - c = \int_0^t [-\bar{v}^2 - p(t_1)\bar{v} - q(t_1)]\, dt_1, \qquad (11.7)$$

где $\bar{v} = \lim_{n\to\infty} v_n(t)$. Следовательно, функция $\bar{v}(t)$ непрерывна, так как она представляется в виде интеграла (11.7). Дифференцируя (11.7), получаем

$$\bar{v}' = -\bar{v}^2 - p(t)\bar{v} - q(t), \quad \bar{v}(0) = c. \qquad (11.8)$$

Из обычной теоремы единственности следует, что $v = \bar{v}$.

Наконец, в силу непрерывности элементов последовательности $\{v_n\}$, ее монотонной сходимости и непрерывности $v(t)$ мы можем утверждать по теореме Дини, что сходимость *равномерна* на $[0, t_0]$.

12. Обсуждение

В предыдущих разделах мы рассмотрели два важных свойства последовательности приближений, получаемой с помощью квазилинеаризации, — *квадратичную сходимость* и *монотонность*. Первое из этих свойств является прямым следствием применения обобщенного метода приближения с помощью касательных — типа метода Ньютона — Рафсона. Второе же свойство справедливо только для специальных классов операторов или для малых интервалов изменения t. В данном случае это есть следствие свойства положительности оператора

$$L = \frac{d}{dt} + p(t), \qquad (12.1)$$

или вернее оператора L^{-1}.

Мы неоднократно использовали тот факт, что

$$Lu \gtrless 0 \Rightarrow u \gtrless 0. \qquad (12.2)$$

К счастью, как мы увидим ниже, многие важные операторы, встречающиеся в анализе и математической физике, обладают этим свойством положительности. В дальнейшем мы будем рассматривать линейные дифференциальные уравнения второго и n-го порядков, параболические и эллиптические дифференциальные уравнения в частных производных и некоторые важные

типы квазилинейных дифференциальных уравнений в частных производных.

Обобщение метода Ньютона — Рафсона на функциональное пространство было сделано Канторовичем и др. Учитывая вклад Канторовича в эту важную область, использованный выше метод приближений можно назвать приложением общего метода Ньютона — Рафсона — Канторовича.

13. Основная лемма

Свойство положительности оператора L^{-1} можно использовать для установления следующего фундаментального результата теории дифференциальных уравнений:

Теорема 2. *Если* $u(t)$, $v(t) \geqslant 0$ *для* $t \geqslant 0$, $c \geqslant 0$, *то из*

$$u(t) \leqslant c + \int_0^t u(s)\, v(s)\, ds \tag{13.1}$$

следует

$$u(t) \leqslant c e^{\int_0^t v(s)\, ds}. \tag{13.2}$$

Этот результат вместе с его обобщениями чрезвычайно полезен для установления существования и единственности решений нелинейных дифференциальных уравнений, а также для исследования устойчивости решений дифференциальных уравнений как в обыкновенных, так и в частных производных. Он принадлежит к общей теории дифференциальных неравенств, основоположником которой является Чаплыгин.

КОММЕНТАРИИ И БИБЛИОГРАФИЯ

2. Детальное обсуждение исторического развития метода Ньютона — Рафсона дано в работе

Whiteside D. T., Patterns of mathematical thought in the latter seventeenth century, *Arch. Hist. of the Exact Sci.*, 1, № 3 (1961), 179—388 (особенно стр. 207).

Об аналитических результатах и обобщениях на функциональное пространство см.

Канторович Л. В., Функциональный анализ и прикладная математика, *УМН*, 3, вып. 6 (1948), 89—185;
Канторович Л. В., Крылов В. И., Приближенные методы высшего анализа, Физматгиз, М.—Л., 1962;
Вайнберг М. М., Вариационные методы исследования нелинейных операторов, Гостехиздат, М., 1956;

Савенко С. С., Об одном итерационном методе решения алгебраических и трансцендентных уравнений, *Журнал вычислительной математики и математической физики*, **4**, № 4 (1964), 738—744, где можно найти много дальнейших ссылок.

4. Минимаксная задача, поставленная в этом разделе, обсуждалась в работах

Sylvester J. J., Meditations on idea of Poncelet's theorem; Notes on meditation ..., *Math. Papers*, 2;
Hammer P. C., *Proceedings Comp. Seminar*, IBM (1949, December), 132.

Насколько нам известно, это единственные опубликованные работы в этой области; однако существуют неопубликованные заметки Беллмана и Тьюки.

5. Этот результат получен в статье

Bellman R., On explicit solutions of some trinomial equations in terms of the maximum operation, *Math. Mag.*, 1956, September — October, стр. 41—44.

6. Теория разрешимости дифференциальных уравнений в элементарных функциях и квадратурах была создана Лиувиллем и в дальнейшем развита Островским и Риттом. Для ознакомления см.

Ritt J. F., Integration in finite terms, Liouville's theory of elementary methods, Columbia University Press, New York, 1948;
Ватсон Г. Н., Теория бесселевых функций, ИЛ, М., 1949.

См. также

Hardy G. H., The integration of functions of a single variable, Cambridge University Press, 1958.

8. Этот результат получен в работе

Bellman R., Functional equations in the theory of dynamic programming — V: Positivity and quasilinearity, *Proc. Nat. Acad. Sci. USA*, **41** (1955), 743—746.

Он также довольно подробно рассмотрен в работе

Kalaba R., On nonlinear differential equations, the maximum operation and monotone convergence, *J. Math. Mech.*, **8** (1959), 519—574.

Явное представление решения уравнения Риккати использовал Калогеро при решении некоторых важных вопросов теории рассеивания; по этому вопросу см.

Calogero F., A novel approach to elementary scattering theory, *Nuovo Cimento*, **27** (1963), 261—302;
Calogero F., A variational principle for scattering phase shifts, *Nuovo Cimento*, **27** (1963), 947—951;
Calogero F., A note on the Riccati equation, *J. Math. Phys.*, **4** (1963), 427—430;
Calogero F., The scattering of a Dirac particle on a central scalar potential, *Nuovo Cimento*, **27** (1963), 1007—1016;
Calogero F., Maximum and minimum principle in potential scattering, *Nuovo Cimento*, **28** (1963), 320—333.

Что касается применения к теории устойчивости, см.

Bellman R., A note on asymptotic behavior of differential equations, *Boll. d'Unione Mate.*, **18** (1963), 16—18.

10—11. Эти результаты впервые были получены в работе Р. Калабы, упомянутой выше. О применении к стохастическим дифференциальным уравнениям см.

Bellman R., On the representation of the solution of a class of stochastic differential equations, *Proc. Amer. Math. Soc.*, **9** (1958), 326—327.

Что касается обобщений предыдущих результатов на случай уравнения $u'=g(u, t)$, см.

Bellman R., On monotone convergence to solutions of $u'=g(u, t)$, *Proc. Amer. Math. Soc.*, **8** (1957), 1007—1009.

Очень интересное новое применение методов приближения второго порядка изложено в работе

Moser J., A new technique for the construction of solutions of nonlinear differential equations, *Proc. Nat. Acad. Sci. USA*, **47** (1961), 1824—1831.

13. Этот результат был впервые использован для исследования устойчивости в работах

Bellman R., The boundedness of solutions of linear differential equations, *Duke Math. J.*, **14** (1947), 83—97;

Беллман Р., Теория устойчивости решений дифференциальных уравнений, ИЛ, М., 1964.

Обобщения этой леммы получили Бихари и Лангенхоп; ссылки на них и на С. А. Чаплыгина см. в книге

Беккенбах Э., Беллман Р., Неравенства, изд-во «Мир», М., 1965.

О применении этой леммы к исследованию теорем существования и единственности решений обыкновенных дифференциальных уравнений см.

Немыцкий В. В., Степанов В. В., Качественная теория дифференциальных уравнений, Физматгиз, М., 1949.

Что касается общего обзора нелинейных задач, см.

T. von Kárman, The engineer grapples with nonlinear problems, *Bull. Amer. Math. Soc.*, **46** (1940), 615—683;

Temple G., Linearization and delinearization, *Proc. International Congress of Math.*, 1958, 233—247.

С некоторыми приложениями читатель может ознакомиться в работах

Radbill J. R., Application of quasilinearization to boundary-layer equations, *AIAA Journal*, **2**, № 10 (1964), 8—10;

Kumar K. S. P., Sridhar R., On the identification of control systems by the quasilinearization method, *IEEE Trans. on Automatic Control*, **AC-9** (1964, April), 151—154;

Smith M. C., Wu C. S., Schwimmer H. S., Magnetohydrodynamic-hypersonic viscous and inviscid flow near the stagnation point of a blunt body, The RAND Corporation, RM-4376-PR, 1965, January;

Kalaba R., Dynamic programming, Fermat's principle and the Eikonal equation, *J. Optical Soc. Amer.*, **51** (1961, October), 1150—1151.

Глава II

Двухточечные краевые задачи для дифференциальных уравнений второго порядка

1. Введение

Приступим теперь к изучению нелинейных дифференциальных уравнений второго порядка вида

$$u'' = g(u, u', t) \tag{1.1}$$

с двухточечными краевыми условиями $u(0) = a_1$, $u(b) = a_2$. Мы не располагаем таким удобным и эффективным способом для выражения общего решения посредством конечного набора частных решений, как в линейном случае (вопрос, который мы обсудим ниже). Следовательно, мы не располагаем готовым рецептом для того, чтобы свести задачу, возникающую при решении уравнения (1.1), к алгебраической задаче, как это имеет место, когда функция $g(u, u', t)$ линейна относительно u и u'. В самом деле, вопросы, связанные с существованием и единственностью решений нелинейных дифференциальных уравнений, весьма сложны и тонки, и в этой области есть еще много белых пятен.

Для того чтобы построить аналитический фундамент и в то же время разработать вычислительные алгоритмы, мы вынуждены прибегнуть к методике аппроксимаций. Методы неподвижной точки, столь эффективные при доказательстве существования решения, бесполезны с вычислительной точки зрения. Вообще лишь немногие стандартные классические методы последовательных приближений могут быть использованы в вычислительной практике. Тем не менее в настоящее время существует ряд мощных вычислительных методов. Мы остановимся на одном частном методе, который вытекает из теории квазилинеаризации.

Наша цель — получить решение уравнения (1.1), если оно существует, как предел последовательности решений *линейных* дифференциальных уравнений. Каждое из этих уравнений нетрудно решить численно, пользуясь фундаментальным принципом суперпозиции. Кроме того, так же как и в случае уравнения Риккати, которое рассматривалось в предыдущей главе, мы хотим обеспечить квадратичную сходимость и, если это возможно, монотонность. Результаты, которые мы получим, покажут, что эти желания реальны.

2. Двухточечные краевые задачи для линейного дифференциального уравнения второго порядка

Прежде чем перейти к изучению общего дифференциального уравнения второго порядка, рассмотрим двухточечную краевую задачу для линейного уравнения

$$u'' + p(t)u' + q(t)u = 0, \quad u(0) = a_1, \quad u(b) = a_2. \tag{2.1}$$

Мы намеренно выбрали особенно простые краевые условия, так как для любого вида краевых условий подход остается одним и тем же.

Пусть функции u_1 и u_2, образующие фундаментальную систему решений записанного выше однородного уравнения, удовлетворяют начальным условиям

$$\begin{aligned} u_1(0) &= 1, \quad u_2(0) = 0, \\ u_1'(0) &= 0, \quad u_2'(0) = 1. \end{aligned} \tag{2.2}$$

Тогда в силу линейности уравнения (2.1) любое решение можно представить как линейную комбинацию

$$u = c_1u_1 + c_2u_2, \tag{2.3}$$

где константы c_1 и c_2 определяются из начальных или краевых условий. Используя условия при $t=0$ и $t=b$, мы без труда получаем два линейных алгебраических уравнения

$$\begin{aligned} a_1 &= c_1, \\ a_2 &= c_1u_1(b) + c_2u_2(b). \end{aligned} \tag{2.4}$$

отсюда

$$\begin{aligned} c_1 &= a_1, \\ c_2 &= \frac{a_2 - a_1u_1(b)}{u_2(b)}. \end{aligned} \tag{2.5}$$

Таким образом, если $u_2(b) \neq 0$, то существует единственное решение. Если $u_2(b) = 0$, то решение существует при $a_2 = a_1u_1(b)$, и в этом случае мы имеем однопараметрическое семейство решений вида

$$u(t) = a_1u_1(t) + c_2u_2(t), \tag{2.6}$$

где c_2 — произвольная постоянная. Следовательно, существование и единственность здесь тесно взаимосвязаны, как это часто и бывает.

Условие $u_2(b) \neq 0$ можно рассматривать как условие на собственные значения задачи Штурма — Лиувилля

$$\begin{aligned} u'' + p(t)u' + \lambda q(t)u &= 0, \\ u(0) = u(b) &= 0. \end{aligned} \tag{2.7}$$

Условие $u_2(b) \neq 0$ эквивалентно требованию, чтобы $\lambda = 1$ не являлось собственным значением. Другими словами, мы требуем, чтобы в случае $\lambda = 1$ единственное решение уравнения (2.7) было тривиальным решением $u(t) = 0$.

3. Неоднородное уравнение

Займемся теперь неоднородным уравнением

$$\begin{gathered} u'' + p(t)\,u' + q(t)\,u = r(t), \\ u(0) = a_1, \quad u(b) = a_2. \end{gathered} \tag{3.1}$$

Воспользовавшись линейностью уравнения, запишем

$$u = v + w, \tag{3.2}$$

где w и v выбраны таким образом, чтобы удовлетворять уравнениям

$$\begin{gathered} w'' + p(t)\,w' + q(t)\,w = r(t), \\ w(0) = 0, \quad w'(0) = 0, \end{gathered} \tag{3.3}$$

и

$$\begin{gathered} v'' + p(t)\,v' + q(t)\,v = 0, \\ v(0) = a_1, \quad v(b) = a_2 - w(b). \end{gathered} \tag{3.4}$$

Для первого уравнения поставлена задача Коши, которую мы рассмотрим ниже; решение краевой задачи для второго уравнения уже нами обсуждалось.

Остается исследовать аналитическое решение неоднородного уравнения (3.3) с помощью решений однородного уравнения (2.1). Наиболее изящный и эффективный способ связан с применением теории матриц; этот способ мы кратко изложим в следующем разделе.

Однако сейчас, чтобы не нарушать изложение, используем простой метод Лагранжа вариации произвольных постоянных. Пусть u_1 и u_2 — частные решения однородного уравнения, которые упоминались выше. Составим их линейную комбинацию

$$w = s_1 u_1 + s_2 u_2, \tag{3.5}$$

где s_1 и s_2 — функции от t, которые мы определим далее. Тогда

$$w' = s_1 u_1' + s_2 u_2' + s_1' u_1 + s_2' u_2. \tag{3.6}$$

В целях упрощения положим

$$s_1' u_1 + s_2' u_2 = 0 \tag{3.7}$$

Это первое условие на s_1' и s_2'. Так как при этом ограничении $w' = s_1 u_1' + s_2 u_2'$, то

$$w'' = s_1 u_1'' + s_2 u_2'' + s_1' u_1' + s_2' u_2'. \tag{3.8}$$

Таким образом,

$$w'' + p(t)\, w' + q(t)\, w = s_1 (u_1'' + p(t)\, u_1' + q(t)\, u_1) + \\ + s_2 (u_2'' + p(t)\, u_2' + q(t)\, u_2) + s_1' u_1' + s_2' u_2' = r(t). \tag{3.9}$$

Поскольку u_1 и u_2 удовлетворяют однородному уравнению, отсюда следует, что

$$s_1' u_1' + s_2' u_2' = r(t). \tag{3.10}$$

Объединяя это уравнение и уравнение (3.7), получаем систему линейных алгебраических уравнений относительно s_1' и s_2'. Существование решения обеспечивается отличием от нуля определителя (вронскиана)

$$W(t) = \begin{vmatrix} u_1 & u_2 \\ u_1' & u_2' \end{vmatrix}. \tag{3.11}$$

Принимая этот факт без доказательства (оно будет приведено в гл. III, разд. 3), имеем

$$s_1' = \frac{\begin{vmatrix} 0 & u_2 \\ r(t) & u_2' \end{vmatrix}}{W(t)}, \quad s_2' = \frac{\begin{vmatrix} u_1 & 0 \\ u_1' & r(t) \end{vmatrix}}{W(t)}. \tag{3.12}$$

Так как мы стремимся удовлетворить условиям $w(0) = w'(0) = 0$, положим $s_1(0) = s_2(0) = 0$, и тогда

$$s_1 = -\int_0^t \frac{u_2(t_1)\, r(t_1)\, dt_1}{W(t_1)}, \\ s_2 = \int_0^t \frac{u_1(t_1)\, r(t_1)\, dt_1}{W(t_1)}. \tag{3.13}$$

Таким образом,

$$w = \int_0^t r(t_1) \frac{[u_1(t_1)\, u_2(t) - u_2(t_1)\, u_1(t)]\, dt_1}{W(t_1)} = \int_0^t G(t, t_1)\, r(t_1)\, dt_1. \tag{3.14}$$

Для того чтобы решить краевую задачу

$$u'' + p(t)\, u' + q(t)\, u = r(t), \quad u(0) = u(b) = 0, \tag{3.15}$$

поступим следующим образом. Решение уравнения (3.15) при условии $u(0)=0$ имеет вид

$$u = c_2 u_2(t) + \int_0^t G(t, t_1) r(t_1)\, dt_1, \tag{3.16}$$

где c_2 — константа, которую нужно определить. Положив $t=b$, получим

$$c_2 = -\int_0^b \frac{G(b, t_1) r(t_1)\, dt_1}{u_2(b)}. \tag{3.17}$$

Следовательно,

$$u = \frac{-u_2(t)}{u_2(b)} \int_0^b G(b, t_1) r(t_1)\, dt_1 + \int_0^t G(t, t_1) r(t_1)\, dt_1 =$$

$$= \frac{-u_2(t)}{u_2(b)} \int_t^b G(b, t_1) r(t_1)\, dt_1 +$$

$$+ \int_0^t \left[G(t, t_1) - \frac{u_2(t)}{u_2(b)} G(b, t_1) \right] r(t_1)\, dt_1. \tag{3.18}$$

Простая выкладка показывает, что можно записать

$$u(t) = \int_0^b K(t, t_1, b) r(t_1)\, dt_1, \tag{3.19}$$

где

$$K(t, t_1, b) = \begin{cases} G(t, t_1) - \dfrac{u_2(t)}{u_2(b)} G(b, t_1), & 0 \leqslant t_1 \leqslant t, \\ -\dfrac{u_2(t)}{u_2(b)} G(b, t_1), & t \leqslant t_1 \leqslant b. \end{cases} \tag{3.20}$$

4. Матричный подход

Отметим, между прочим, что прямой вывод формулы (3.14) можно получить применением элементов теории матриц к линейным дифференциальным уравнениям.

Начнем с замечания, что уравнение

$$u'' + p(t) u' + q(t) u = r(t) \tag{4.1}$$

можно записать в виде системы линейных уравнений первого порядка

$$\begin{aligned} u_1' &= u_2, \\ u_2' &= -p(t) u_2 - q(t) u_1 + r(t). \end{aligned} \tag{4.2}$$

Отсюда, вводя векторы

$$x(t)=\begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix}, \quad f(t)=\begin{bmatrix} 0 \\ r(t) \end{bmatrix}$$

и матрицу

$$A(t)=\begin{bmatrix} 0 & 1 \\ -q(t) & -p(t) \end{bmatrix}, \tag{4.3}$$

мы получаем векторную систему

$$x'=A(t)x+f(t), \tag{4.4}$$

эквивалентную системе (4.2), с начальным условием $x(0)=0$.

Существуют два подхода для получения решения системы (4.4). В первом используется метод Лагранжа вариации произвольных постоянных в матричном виде. Во втором подходе, который можно обобщить на любой тип линейных функциональных уравнений, используется решение сопряженного уравнения. Остановимся на втором подходе.

Пусть $Y(t)$ — пока произвольная матрица. Составим равенство

$$\int_0^T Yx'\,dt=\int_0^T [YA(t)x+Yf(t)]\,dt. \tag{4.5}$$

Интегрируя по частям, получаем

$$Y(T)x(T)-Y(0)x(0)-\int_0^T Y'x\,dt= \\ =\int_0^T YA(t)x\,dt+\int_0^T Yf(t)\,dt. \tag{4.6}$$

Так как $x(0)=0$, это выражение можно переписать в виде

$$Y(T)x(T)=\int_0^T [Y'+YA(t)]\,x\,dt+\int_0^T Yf(t)\,dt. \tag{4.7}$$

Поскольку это соотношение справедливо при произвольной матрице Y, постараемся выбрать ее наиболее удобным образом. Определим матрицу $Y(t)=Y(t, T)$ на отрезке $[0, T]$ уравнением

$$Y'+YA(t)=0, \quad Y(T)=I. \tag{4.8}$$

При таком выборе $Y(t)$ имеем

$$x(T)=\int_0^T Yf(t)\,dt. \tag{4.9}$$

Если обозначить через $X(t)$ решение матричного уравнения

$$X' = A(t)X, \quad X(0) = I, \tag{4.10}$$

то легко видеть, что решение уравнения (4.8) записывается в виде

$$Y = X(T)X(t)^{-1}. \tag{4.11}$$

Это следует из соотношения

$$\frac{d}{dt}(X^{-1}) = -X^{-1}\frac{dX}{dt}X^{-1}. \tag{4.12}$$

Тогда (4.9) принимает вид

$$x(T) = \int_0^T X(T)X(t)^{-1} f(t)\,dt. \tag{4.13}$$

Произведя указанные операции над матрицами, мы получим результат, в точности совпадающий с (3.14).

Соотношение (4.13) более предпочтительно для аналитических целей, поскольку в такой записи весьма наглядно проявляется структура решения.

5. Функции Грина

Весьма важная функция $K(t, t_1, b)$, введенная в разд. 3, носит название функции Грина. Она определяется как видом конкретного линейного уравнения, так и граничными условиями. Например, если уравнение имеет вид

$$u'' = r(t), \quad u(0) = u(b) = 0, \tag{5.1}$$

то простые вычисления по указанной выше схеме приводят к решению

$$u = \int_0^b K(t, t_1, b)r(t_1)\,dt_1, \tag{5.2}$$

где

$$K(t, t_1, b) = \begin{cases} \dfrac{t(t_1 - b)}{b}, & t \leqslant t_1 \leqslant b, \\ \dfrac{t_1(t - b)}{b}, & 0 \leqslant t_1 \leqslant t. \end{cases} \tag{5.3}$$

Ниже мы используем этот результат.

Существует много интересных и важных свойств функций Грина, на которых мы не будем здесь останавливаться. Эти

функции играют основную роль в современной теории дифференциальных уравнений и связанных с ними интегральных уравнений. В следующей главе мы обсудим некоторые важные вопросы о положительности функций Грина.

6. Выпуклость

Обобщим теперь путем введения понятия выпуклости представление

$$u^2 = \max_v (2uv - v^2), \tag{6.1}$$

использованное нами в гл. I для изучения уравнения Риккати. Говорят, что функция $f(x)$ выпукла на отрезке $a \leqslant x \leqslant b$, если $f''(x) \geqslant 0$ на этом отрезке, и строго выпукла, если $f''(x) > 0$.

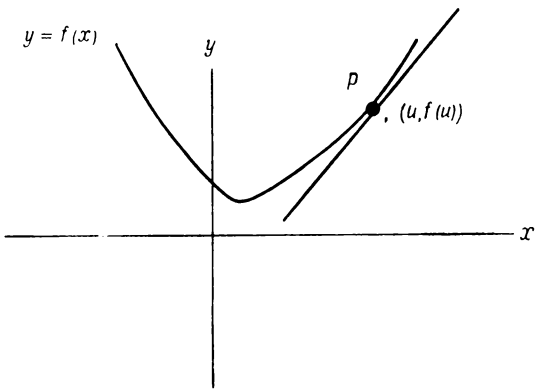


Рис. 2.1

Рассмотрим график строго выпуклой функции (рис. 2.1) и проведем к ней касательную в точке $P(u, f(u))$:

$$y - f(u) = f'(u)(x - u). \tag{6.2}$$

В силу строгой выпуклости кривая расположена выше касательной при всех значениях x, кроме $x = u$. Отсюда

$$f(x) \geqslant f(u) + f'(u)(x - u) \tag{6.3}$$

для всех u и x; равенство в (6.3) имеет место при $x = u$. Поэтому можно написать

$$f(x) = \max_u [f(u) + (x - u) f'(u)]. \tag{6.4}$$

Разумеется, этот результат можно получить чисто аналитическим путем.

Точно так же для функций двух переменных

$$f(x, y) = \max_{u, v} [f(u, v) + (x - u) f_u + (y - v) f_v] \tag{6.5}$$

при условии, что $f(x, y)$ строго выпукла в интересующей нас области, т. е. если

$$f_{xx} > 0, \quad \begin{vmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{vmatrix} > 0 \tag{6.6}$$

или, что то же самое, если квадратичная форма

$$f_{xx} r^2 + 2 f_{xy} rs + f_{yy} s^2 \tag{6.7}$$

положительно определена.

Аналогичные результаты справедливы для функций произвольного конечного числа переменных и для функционалов. Общее исследование выпуклости связано с некоторыми трудностями, однако оно вознаграждается, так как приложения этого простого, но фундаментального понятия всюду встречаются в анализе и геометрии.

7. Квазилинеаризация

Итак, мы уже готовы к применению квазилинеаризации в аналитических и вычислительных целях. Сначала мы используем эту методику для получения квадратичной сходимости и проиллюстрируем это на нескольких примерах.

Для иллюстрации начнем с уравнения

$$u'' = f(u), \quad u(0) = u(b) = 0. \tag{7.1}$$

Пусть $u_0(x)$ — некоторое начальное приближение; рассмотрим последовательность $\{u_n\}$, определяемую рекуррентным соотношением

$$\begin{gathered} u_n'' = f(u_{n-1}) + (u_n - u_{n-1}) f'(u_{n-1}), \\ u_n(0) = u_n(b) = 0. \end{gathered} \tag{7.2}$$

Каждая функция $u_n(x)$ есть решение линейного уравнения, что является весьма важной характерной особенностью этого алгоритма.

Алгоритм следует из метода аппроксимации Ньютона — Рафсона — Канторовича в функциональном пространстве. На основании всего ранее сказанного можно ожидать квадратичную сходимость, если сходимость вообще имеет место.

8. Обсуждение

Поскольку каждое решение определяется двухточечными краевыми условиями, заранее отнюдь не очевидно, что последовательность $\{u_n\}$, задаваемая соотношением (7.2), действительно существует.

Поэтому мы вначале покажем, что упомянутая последовательность полностью определена, если область изменения x ограничена достаточно малым отрезком $[0, b]$, а затем докажем квадратичную сходимость последовательности. В следующей главе мы приведем доказательство монотонной сходимости, что потребует привлечения несколько более сложных идей, и используем этот результат для расширения области сходимости аппроксимирующей последовательности. Мы продемонстрируем ряд различных подходов к решению весьма важных и интересных вопросов, связанных с монотонностью и положительностью.

9. Существование и ограниченность

Прежде всего докажем существование и равномерную ограниченность последовательности $\{u_n(x)\}$ при достаточно малом b. Исходя из уравнения (7.2), точнее

$$\begin{gathered} u''_{n+1} = f(u_n) + (u_{n+1} - u_n) f'(u_n), \\ u_{n+1}(0) = u_{n+1}(b) = 0, \end{gathered} \tag{9.1}$$

получим, как показано в разд. 4 и 5, линейное интегральное уравнение

$$u_{n+1} = \int_0^b K(x, y) [f(u_n) + (u_{n+1} - u_n) f'(u_n)] \, dy, \tag{9.2}$$

где $K(x, y)$ — функция Грина,

$$K(x, y) = \begin{cases} \dfrac{x(y-b)}{b}, & 0 \leqslant x < y \leqslant b, \\ \dfrac{(x-b)y}{b}, & b \geqslant x > y \geqslant 0. \end{cases} \tag{9.3}$$

Для того чтобы получить (9.2), представим (9.1) в виде уравнения $u'' = r(t)$, $u(0) = u(b) = 0$, и затем используем результат, установленный в разд. 5.

Заметим, что

$$\max_{x, y} |K(x, y)| = \frac{b}{4}, \tag{9.4}$$

где максимизация проведена по области $0 \leqslant x,\ y \leqslant b$. Положим

$$\max_{|u| \leqslant 1} \max(|f(u)|,\ |f'(u)|) = m, \tag{9.5}$$

предполагая $m < \infty$, и выберем $u_0(x)$ так, чтобы $|u_0(x)| \leqslant 1$ при $0 \leqslant x \leqslant b$. Возвращаясь к (9.2), имеем

$$|u_{n+1}| \leqslant \int_0^b |K(x, y)|\,[|f(u_n)| + |u_n|\,|f'(u_n)| + |f'(u_n)|\,|u_{n+1}|]\,dy. \tag{9.6}$$

Отсюда, обозначая $m_1 = \max_{0 \leqslant x \leqslant b} |u_1(x)|$, получаем при $n=0$

$$m_1 \leqslant \frac{b}{4} \int_0^b [2m + mm_1]\,dy \leqslant \frac{b^2 m}{2} + \frac{b^2 m}{4} m_1. \tag{9.7}$$

Следовательно, при условии $b^2 m/4 < 1$, находим границу

$$m_1 \leqslant \frac{b^2 m/2}{1 - b^2 m/4}. \tag{9.8}$$

Эта верхняя граница m_1 не будет превосходить 1, если выполнено неравенство $b^2 \leqslant 4/3m$. Поэтому, если выбрать отрезок $[0, b]$ достаточно малым, так, чтобы выполнялось условие $b^2 \leqslant 4/3m$, будем иметь $m_1 \leqslant 1$.

Ясно, что эти рассуждения можно продолжить по индукции, и в результате можно утверждать, что $|u_n(x)| \leqslant 1$ при $0 \leqslant x \leqslant b$, если $b^2 \leqslant 4/3m$. Итак, мы показали, что принятое нами индуктивное определение последовательности $\{u_n(x)\}$ имеет смысл.

10. Сходимость

Доказательство сходимости будет проведено стандартным путем; единственное отличие будет состоять в том, что сходимость окажется квадратичной.

Обращаясь к рекуррентному соотношению (9.1), вычтем n-е уравнение из $(n+1)$-го:

$$(u_{n+1} - u_n)'' = f(u_n) - f(u_{n-1}) - (u_n - u_{n-1}) f'(u_{n-1}) + f'(u_n)(u_{n+1} - u_n). \tag{10.1}$$

Рассматривая (10.1) как уравнение относительно $u_{n+1} - u_n$ и преобразуя его в интегральное уравнение, как это уже делалось, получаем

$$(u_{n+1} - u_n) = \int_0^b K(x, y)\,\{f(u_n) - f(u_{n-1}) - (u_n - u_{n-1}) f'(u_{n-1}) + f'(u_n)(u_{n+1} - u_n)\}\,dy. \tag{10.2}$$

Из теоремы о среднем следует

$$f(u_n) - f(u_{n-1}) - (u_n - u_{n-1}) f'(u_{n-1}) = \frac{(u_n - u_{n-1})^2}{2} f''(\theta), \quad (10.3)$$

где θ лежит между u_{n-1} и u_n. Тогда, полагая

$$k = \max_{|u| \leqslant 1} |f''(u)|, \quad (10.4)$$

имеем почти как раньше

$$|u_{n+1} - u_n| \leqslant \frac{b}{4} \int_0^b \left[\frac{k}{2}(u_n - u_{n-1})^2 + m|u_{n+1} - u_n|\right] dy. \quad (10.5)$$

Отсюда, проводя рассуждения, аналогичные тем, которые использовались при выводе неравенств (9.7) и (9.8), получаем

$$\max_x |u_{n+1} - u_n| \leqslant \frac{b^2 k/8}{1 - b^2 m/4} \left(\max_x |u_n - u_{n-1}|\right)^2. \quad (10.6)$$

Это соотношение показывает, что если сходимость вообще имеет место, то она квадратичная.

Если члены последовательности $\{v_n\}$, $n = 0, 1, 2, \ldots$, удовлетворяют соотношению

$$v_n \leqslant k_1 v_{n-1}^2, \quad n = 1, 2, \ldots, \quad (10.7)$$

то обычная индукция приводит к неравенству

$$v_n \leqslant k_1^{2^n - 1} v_0^{2^n} = \frac{(k_1 v_0)^{2^n}}{k_1}. \quad (10.8)$$

Таким образом, сходимость будет зависеть от величины

$$k_1 v_0 = \frac{b^2 k/8}{1 - b^2 m/4} \max_x |u_1 - u_0|, \quad (10.9)$$

которая (это можно показать таким же путем, как и выше) будет меньше единицы при достаточно малом b.

Интересно заметить, что, как это следует из доказательства, для сходимости достаточно малости $\max_x |u_{n+1} - u_n|$ хотя бы при *одном* n. Следовательно, даже если заданный отрезок $[0, b]$ окажется слишком большим, можно всегда надеяться, что за счет удачного выбора начального приближения $u_0(x)$ мы добьемся достаточно малой величины $|u_1(x) - u_0(x)|$. Здесь большую помощь может оказать физический и инженерный опыт, а также интуиция.

Итак, если величина (10.9) меньше единицы, то последовательность $\{u_n(x)\}$ равномерно сходится к функции $u(x)$, удовле-

творяющей уравнению

$$u = \int_0^b K(x, y) f(u)\, dy. \tag{10.10}$$

Отсюда мы заключаем, что u удовлетворяет исходному дифференциальному уравнению.

Кроме того, пользуясь вышеизложенными методами, можно показать, что

$$\max_x |u - u_n| \leqslant k_2 \max_x |u - u_{n-1}|^2, \tag{10.11}$$

откуда, как обычно, следует единственность.

Мы тщательно вникли во все детали доказательства существования и сходимости последовательности приближений, чтобы продемонстрировать необходимые этапы, из которых должно состоять доказательство. Поэтому при исследовании более сложных функциональных уравнений в последующих главах мы будем опускать доказательства существования и сходимости, поскольку их проведение не предполагает привлечения дополнительных идей или методов — достаточно самого общего знакомства с обозначениями теории матриц и функционального анализа. Заметим, что предположение о выпуклости, столь существенное при доказательстве теоремы представления, не играет никакой роли при доказательстве квадратичной сходимости.

11. Сходимость алгоритма Пикара

Обычный подход, основанный на рекуррентном соотношении

$$v''_{n+1} = f(v_n), \quad v_{n+1}(0) = v_{n+1}(b) = 0, \tag{11.1}$$

при почти тех же ограничениях на f и b приводит к сходящейся последовательности. Однако в общем случае скорость сходимости будет такой же, как для геометрической прогрессии, иначе говоря,

$$|v_{n+1} - v_n| \leqslant k |v_n - v_{n-1}|, \tag{11.2}$$

где $k<1$, или $|v_n - v_{n-1}| \leqslant k^n c_0$.

Преимущество квадратичной сходимости состоит, разумеется, в быстроте сходимости. Кроме того, мы покажем в дальнейшем (в гл. III), что квазилинеаризация в ряде важных случаев обеспечивает максимальный интервал сходимости последовательных приближений.

12. Числовой пример

В магнитной гидродинамике в настоящее время проявляется интерес к уравнению

$$u_{xx} + u_{yy} = e^u. \tag{12.1}$$

Его численное решение с использованием квазилинеаризации будет рассмотрено в дальнейшем. Здесь мы обсудим его одномерный вариант

$$u'' = e^u, \quad u(0) = u(b) = 0, \tag{12.2}$$

существенное достоинство которого состоит в том, что оно имеет единственное и явное аналитическое решение. При $b = 1$ решение имеет вид

$$u(x) = -\ln 2 + 2\ln\left[c \sec\left\{\frac{c\,(x - 1/2)}{2}\right\}\right], \tag{12.3}$$

где c — корень уравнения

$$\sqrt{2} = c \sec\left(\frac{c}{4}\right), \tag{12.4}$$

лежащий между 0 и $\pi/2$, а именно с точностью до восьмого знака $c = 1{,}3360557$.

Напомним, как получено решение уравнения (12.2). Умножая на u' обе части уравнения и интегрируя, мы получаем u' как функцию от u. Дальнейшее интегрирование приводит к искомому явному решению уравнения (12.3).

Используя методы, изложенные выше, получаем аппроксимирующую последовательность

$$u''_{n+1} = e^{u_n} + e^{u_n}(u_{n+1} - u_n), \quad u_{n+1}(0) = u_{n+1}(b) = 0. \tag{12.5}$$

Таблица 2.1

Приближение $u_n(x)$ к $u(x)$

x	$u_0(x)$	$u_1(x)$	$u_2(x)$	$u(x)$
0,0	0,000000	0,000000	0,000000	0,000000
0,1	0,000000	—0,*041*285	—0,*041436*	—0,041436
0,2	0,000000	—0,*072*974	—0,*0732683*	—0,0732686
0,3	0,000000	—0,*095*386	—0,*095800*	—0,095800
0,4	0,000000	—0,*108*743	—0,*109238*	—0,109238
0,5	0,000000	—0,*113*181	—0,*113704*	—0,113704
0,6	0,000000	—0,*108*743	—0,*109238*	—0,109238
0,7	0,000000	—0,*095*386	—0,*095800*	—0,095800
0,8	0,000000	—0,*072*974	—0,*0732683*	—0,0732686
0,9	0,000000	—0,*041*285	—0,*041436*	—0,041436
1,0	0,000000	0,000000	0,000000	0,000000

Выбирая наиболее естественное начальное приближение $u_0(x)=0$, мы вычисляем функции $u_1(x)$ и $u_2(x)$ так, как это описано выше, используя метод интегрирования Рунге — Кутта. Табл. 2.1 иллюстрирует скорость сходимости. Значащие цифры величины $u_n(x)$, совпадающие со знаками $u(x)$, выделены *курсивом*.

13. Нелинейные дифференциальные уравнения второго порядка общего вида

Нетрудно применить тот же метод для исследования уравнения

$$u''=f(u',\ u,\ x) \tag{13.1}$$

с нелинейными краевыми условиями вида

$$g_1(u(0),\ u'(0))=0,\quad g_2(u(b),\ u'(b))=0 \tag{13.2}$$

или даже более общего вида

$$\begin{aligned} &g_1(u(0),\ u'(0),\ u(b),\ u'(b))=0,\\ &g_2(u(0),\ u'(0),\ u(b),\ u'(b))=0. \end{aligned} \tag{13.3}$$

Теперь мы применим квазилинеаризацию как к уравнению, так и к граничным условиям. Тогда в случае уравнения (13.1) и условий (13.2) мы получаем последовательность $\{u_n(x)\}$ при помощи уравнения

$$\begin{aligned} u''_{n+1}=f(u'_n,\ u_n,\ x)+f_{u'}(u'_n,\ u_n,\ x)(u'_{n+1}-u'_n)+\\ +f_u(u'_n,\ u_n,\ x)(u_{n+1}-u_n) \end{aligned} \tag{13.4}$$

с линеаризованными краевыми условиями

$$\begin{aligned} g_1(u_n(0),\ u'_n(0))+g_{1u}(u_n(0),\ u'_n(0))(u_{n+1}(0)-u_n(0))+\\ +g_{1u'}(u_n(0),\ u'_n(0))(u'_{n+1}(0)-u'_n(0))=0. \end{aligned} \tag{13.5}$$

Аналогичное соотношение выписывается для g_2 в точке $x=b$.

Легко доказать существование, единственность и квадратичную сходимость при достаточно малом b и очевидных ограничениях на величины f_u, $f_{u'}$, g_{iu}, $g_{iu'}$, $i=1,\ 2$.

14. Числовой пример

В качестве второго примера рассмотрим уравнение

$$-u''=1+a^2(u')^2,\quad u(0)=u(b)=0, \tag{14.1}$$

возникающее при изучении конечных деформаций упругой струны под действием поперечной нагрузки[1]. Точное аналитическое решение при $a^2=0{,}49$ и $b=1$ имеет вид

$$u(x)=\frac{1}{a^2}\ln\left[\frac{\cos a\,(x-1/2)}{\cos a/2}\right]. \tag{14.2}$$

[1]) Это уравнение предложил нам В. Прагер.

Используя рекуррентное соотношение

$$-u''_{n+1} = 1 - 0{,}49\left(u'^2_n\right) + 2\,(0{,}49)\,u'_n u'_{n+1},$$
$$u_{n+1}(0) = u_{n+1}(b) = 0, \tag{14.3}$$

с начальным приближением $u_0(x) = 0$, мы получаем результаты, которые приведены в табл. 2.2.

Таблица 2.2

Сходимость $u_n(x)$ к $u(x)$

x	$u_0(x)$	$u_1(x)$	$u_2(x)$	$u(x)$
0,0	0,000000	0,000000	0,000000	0,000000
0,1	0,000000	0,045000	0,046570	0,046571
0,2	0,000000	0,080000	0,082302	0,082304
0,3	0,000000	0,105000	0,107571	0,107573
0,4	0,000000	0,120000	0,122632	0,122635
0,5	0,000000	0,125000	0,127636	0,127639
0,6	0,000000	0,120000	0,122632	0,122635
0,7	0,000000	0,105000	0,107571	0,107573
0,8	0,000000	0,080000	0,082302	0,082304
0,9	0,000000	0,045000	0,046570	0,046571
1,0	0,000000	0,000000	0,000000	0,000000

15. Вариационное исчисление

Естественным и наиболее обильным источником нелинейных дифференциальных уравнений с двухточечными краевыми условиями является вариационное исчисление. Рассмотрим задачу максимизации функционала

$$J(u) = \int_0^b g(u, u')\,dt \tag{15.1}$$

с начальным условием $u(0) = c$. Обращение в нуль первой вариации приводит к уравнению Эйлера

$$\frac{\partial g}{\partial u} - \frac{d}{dt}\left(\frac{\partial g}{\partial u'}\right) = 0 \tag{15.2}$$

с конечным условием [1])

$$\left.\frac{\partial g}{\partial u'}\right|_{t=b} = 0. \tag{15.3}$$

[1]) Условие (15.3) — это условие трансверсальности для рассматриваемой вариационной задачи (см., например, Гельфанд И. М. и Фомин С. В., Вариационное исчисление, Физматгиз, М., 1961). — *Прим. ред.*

Если функция g представлена в виде суммы квадратичной формы относительно u и u' и линейного члена относительно u, то уравнение Эйлера линейно и двухточечная краевая задача может быть непосредственно решена. В общем случае перед нами предстают обычные трудности.

Чтобы проиллюстрировать метод квазилинеаризации, рассмотрим задачу, которая возникает при использовании принципа Ферма. Пусть плоскость x, y представляет оптически неоднородную среду; скорость света в точке (x, y) обозначим через $v(x, y)$.

Луч света, исходящий из точки P_0, проходит через точку P_1. Какой путь он избирает?

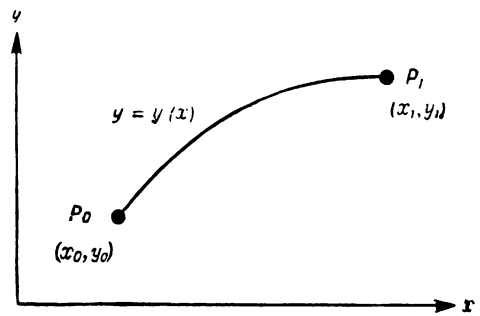


Рис. 2.2. Луч света в неоднородной среде.

Согласно принципу Ферма о минимуме времени, наша задача состоит в минимизации интеграла

$$J(y)=\int_{x_0}^{x_1}\frac{ds}{v(x, y)}=\int_{x_0}^{x_1}\frac{\sqrt{1+y'^2}}{v(x, y)}\,dx \tag{15.4}$$

(рис. 2.2).

Уравнение Эйлера для этой задачи имеет вид

$$\frac{\partial}{\partial y}\left[\frac{\sqrt{1+y'^2}}{v(x, y)}\right]-\frac{d}{dx}\left\{\frac{\partial}{\partial y'}\left[\frac{\sqrt{1+y'^2}}{v(x, y)}\right]\right\}=0. \tag{15.5}$$

Пусть $v(x, y)=y$. Тогда уравнение (15.5) сводится к уравнению

$$1+y'^2+yy''=0 \tag{15.6}$$

с двухточечными краевыми условиями

$$y(x_0)=y_0,\quad y(x_1)=y_1. \tag{15.7}$$

Известно, что оптимальные кривые — это дуги окружностей, центры которых располагаются на оси x. Мы можем использовать этот результат для контроля точности наших численных методов.

Перепишем уравнение (15.6) в виде $y''=-(1+y'^2)/y$ и в качестве начального приближения выберем функцию $z_0(x)$, график которой представляет собой прямую, соединяющую точки P_0 и P_1 .Общее рекуррентное соотношение имеет вид

$$z''_{n+1}=-\frac{\left(1+z'^2_n\right)}{z_n}+\frac{\left(1+z'^2_n\right)}{z_n^2}(z_{n+1}-z_n)-\frac{2z'_n}{z_n}(z'_{n+1}-z'_n), \quad (15.8)$$

где $z_{n+1}(x_0)=y_0$, $z_{n+1}(x_1)=y_1$.

Рассмотрим частный случай, когда $y(1)=1$, $y(2)=2$. В качестве начального приближения выберем $z_0(x)=x$. Полученные численные результаты представлены в табл. 2.3 (в приложении 1 приводится программа).

Таблица 2.3

Численные результаты

x	$z_0(x)$	$z_1(x)$	$z_2(x)$	$z_3(x)$
1,00	1,00	1,000000	1,000000	1,000000
1,25	1,25	1,386195	1,391934	1,391941
1,50	1,50	1,651987	1,658305	1,658312
1,75	1,75	1,848866	1,854642	1,854050
2,00	2,00	2,000000	2,000000	2,000000

Функция $z_3(x)$ совпадает с точным решением до шестого знака включительно.

16. Квазилинеаризация

Использование метода квазилинеаризации может идти двумя путями. С одной стороны, уравнение Эйлера (15.2) является нелинейным дифференциальным уравнением того типа, с которым мы уже встречались. Производя обычную линеаризацию, мы получаем линейное уравнение

$$\begin{aligned} &g_u(u_n, u'_n, t)+g_{uu}(u_n, u'_n, t)(u_{n+1}-u_n)+\\ &\quad+g_{uu'}(u_n, u'_n, t)(u'_{n+1}-u'_n)-\\ &\quad-\frac{d}{dt}\left[g_{u'}(u_n, u'_n, t)+g_{u'u}(u_n, u'_n, t)(u_{n+1}-u_n)+\right.\\ &\qquad\left.+g_{u'u'}(u_n, u'_n, t)(u'_{n+1}-u'_n)\right]=0 \quad (16.1) \end{aligned}$$

с начальным условием $u_{n+1}(0)=c$ и *линеаризованным* конечным условием

$$g_{u'}(u_n, u'_n, b)+(u_{n+1}-u_n)\,g_{u'u}(u_n, u'_n, b)+ \\ +(u'_{n+1}-u'_n)\,g_{u'u'}(u_n, u'_n, b)=0. \quad (16.2)$$

С другой стороны, мы можем действовать непосредственно следующим образом. Пусть u_n — некоторое приближенное решение задачи оптимизации; рассмотрим задачу минимизации функционала

$$J_n=\int_a^b h_2(u, u_n, t)\,dt, \quad (16.3)$$

где $h_2(u, u_n, t)$ — разложение функции $g(u, u', t)$ в окрестности точки (u_n, u'_n), включающее члены второго порядка малости:

$$h_2(u, u_n, t)=g(u_n, u'_n, t)+g_u(u_n, u'_n, t)(u-u_n)+ \\ +g_{u'}(u_n, u'_n, t)(u'-u'_n)+\frac{1}{2}\left[g_{uu}(u_n, u'_n, t)(u-u_n)^2+\right. \\ +2g_{uu'}(u_n, u'_n, t)(u-u_n)(u'-u'_n)+ \\ \left.+g_{u'u'}(u_n, u'_n, t)(u'-u'_n)^2\right]. \quad (16.4)$$

Прямые вычисления показывают, что уравнение Эйлера для функционала J_n в точности совпадает с (16.1). Таким образом, в данном случае приближенный подход к точному вариационному уравнению эквивалентен точному подходу к приближенной вариационной задаче. Одно из преимуществ второго подхода состоит в том, что использование методов динамического программирования и инвариантного погружения позволяет избежать двухточечной краевой задачи, а, следовательно, и необходимости решения линейной системы алгебраических уравнений. Мы подробно обсудим этот вопрос в разд. 20 и гл. VII.

17. Оценки сверху и снизу с помощью квазилинеаризации

Теперь кратко остановимся на том, как, используя квазилинеаризацию, можно получить оценки сверху и снизу для определенного класса вариационных задач. Для иллюстрации рассмотрим задачу минимизации функционала

$$J(u)=\int_0^1 (u'^2+\varphi(u))\,dt, \quad (17.1)$$

при условии, что на u наложены ограничения типа $u(0)=c$. Для простоты возьмем

$$u(0)=c_1, \quad u'(1)=0. \tag{17.2}$$

Второе условие (17.2) следует из того, что на $u(1)$ не наложено ограничений (ср. (15.3)).

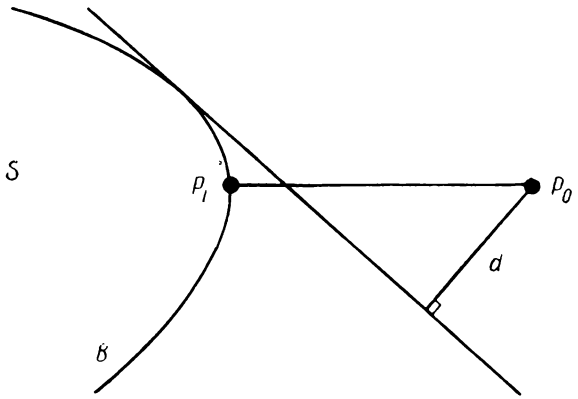


Рис. 2.3.

Предположим, что функция $\varphi(u)$ выпукла, так что можно записать

$$\varphi(u)=\max_v [\varphi(v)+(u-v)\,\varphi'(v)]. \tag{17.3}$$

Тогда

$$\min_u J(u)=\min_u \int_0^1 (u'^2+\varphi(u))\,dt=$$

$$=\min_u \max_v \left[\int_0^1 [u'^2+\varphi(v)+(u-v)\varphi'(v)]\,dt\right]. \tag{17.4}$$

Опираясь на весьма общие результаты теории игр, можно показать, что операции максимизации и минимизации можно поменять местами. Следовательно,

$$\min_u J(u)=\max_v \min_u \left[\int_0^1 (u'^2+\varphi(v)+(u-v)\varphi'(v))\,dt\right]. \tag{17.5}$$

Простейшее доказательство справедливости перестановки основывается на непосредственных вычислениях.

Перестановка совершается для того, чтобы получить нижнюю границу для $J(u)$,

$$\min_u J(u) \geqslant \min_u \left[\int_0^1 (u'^2 + \varphi(v) + (u - v)\varphi'(v))\, dt \right] \quad (17.6)$$

для любой функции $v(t)$. Этот минимум можно вычислить в явном виде, что мы и проделаем ниже.

Прежде чем приступить к вычислению, остановимся немного на двойственности, которую мы использовали при обращении задачи минимизации в задачу максимизации. Рассмотрим общую задачу определения минимального расстояния от точки p_0 до выпуклого множества S (рис. 2.3). Проведем касательные к граничной кривой B и рассмотрим расстояние d от точки p_0 до какой-нибудь из касательных. Геометрически абсолютно ясно, что

$$\overline{p_1 p_0} = \text{минимальное расстояние от } p_0 \text{ до } S = \\ = \text{максимальное расстояние от } p_0 \text{ до касательной} = \max d. \quad (17.7)$$

Кроме того, поскольку d является минимальным расстоянием от точки p_0 до точек касательной, мы сразу же видим, откуда возникает max-min.

18. Решение внутренней задачи

Уравнение Эйлера для задачи минимизации функционала, стоящего в правой части соотношения (17.6), имеет вид

$$u'' - \frac{\varphi'(v)}{2} = 0, \quad u(0) = c, \quad u'(1) = 0. \quad (18.1)$$

Вводя обозначения $u = c + w$, получаем

$$u = c + \frac{1}{2} \int_0^1 k(s, t)\varphi'(v(s))\, ds, \quad (18.2)$$

где $k(s, t)$ — функция Грина для задачи

$$u'' = r, \quad u(0) = u'(1) = 0. \quad (18.3)$$

Простое вычисление показывает, что минимальное значение функционала в правой части неравенства (17.6) равно [1])

$$K(v) = \frac{1}{4} \int_0^1 \int_0^1 k(s, t)\varphi'(v(s))\varphi'(v(t))\, ds\, dt + \\ + \int_0^1 [(c - v)\varphi'(v) + \varphi(v)]\, dt. \quad (18.4)$$

[1]) Чтобы получить (18.4), нужно в интеграле из правой части (17.6) проинтегрировать по частям первый член, а затем воспользоваться соотношениями (18.1) и (18.2). — *Прим. ред.*

Теперь мы можем записать

$$\min_u J(u) = \max_v K(v). \tag{18.5}$$

Левая часть немедленно дает оценки сверху, а правая часть — оценки снизу. Разумный выбор пробных функций u и v часто позволяет добиться довольно узкого интервала оценок.

19. Дальнейшие применения квазилинеаризации

Ту же идею можно использовать при решении некоторых нестандартных задач теории управления и вариационного исчисления. Рассмотрим для примера задачу минимизации неаналитического функционала

$$J(f) = \int_0^1 |1 - u(t)|\, dt, \tag{19.1}$$

определенного для всех функций f, удовлетворяющих ограничениям

$$\begin{gathered} 0 \leqslant f(t) \leqslant M, \quad 0 \leqslant t \leqslant 1, \\ \int_0^1 f\, dt \leqslant a. \end{gathered} \tag{19.2}$$

Кроме того, u и f связаны соотношением

$$u' = -u + f, \quad u(0) = c. \tag{19.3}$$

Записывая

$$|1 - u| = \max_{-1 \leqslant \varphi \leqslant 1} (1 - u)\varphi \tag{19.4}$$

и используя теорему о минимаксе, можно переставить min и max и получить явное аналитическое решение задачи. Подробности можно найти в монографии Беллмана, Гликсберга, Гросса (см. библиографию в конце главы).

20. Динамическое программирование

Коротко остановимся на том, как, используя метод динамического программирования и более общий метод инвариантного погружения, можно избежать решения двухточечной краевой задачи. Изучение этих подходов тем более важно, что основным источником ошибок при использовании квазилинеаризации является решение системы линейных алгебраических уравнений, связанных с двухточечной краевой задачей. В рассмотренных

случаях это не представляет значительной проблемы, однако в задаче высокой размерности может вызвать существенные затруднения.

Для того чтобы продемонстрировать, как динамическое программирование обходит решение краевых задач, рассмотрим задачу минимизации квадратичного функционала

$$J(u)=\int_a^T \left[u'^2+b(t_1)u^2\right]dt_1 \tag{20.1}$$

на классе функций u, для которых существует этот интеграл и которые удовлетворяют начальному условию $u(a)=c$.

Уравнение Эйлера и краевые условия здесь имеют вид

$$u''-b(t)u=0, \quad u(a)=c, \quad u'(T)=0. \tag{20.2}$$

Подойдем, однако, к задаче минимизации по-иному. Запишем

$$\min_u J(u)=f(a, c). \tag{20.3}$$

Тогда принцип оптимальности приводит к нелинейному дифференциальному уравнению в частных производных

$$-\frac{\partial f}{\partial a}=\min_v\left[v^2+b(a)c^2+v\frac{\partial f}{\partial c}\right] \tag{20.4}$$

с начальным условием $f(T, c)=0$. Проведя минимизацию, мы получаем

$$-\frac{\partial f}{\partial a}=b(a)c^2-\frac{1}{4}\left(\frac{\partial f}{\partial c}\right)^2. \tag{20.5}$$

Чтобы решить это уравнение, заметим, что в силу (20.2) u зависит от c линейно и, следовательно, $f(a, c)$ является квадратичной функцией от c:

$$f(a, c)=r(a)c^2. \tag{20.6}$$

Подставляя это выражение в (20.5), видим, что $r(a)$ удовлетворяет уравнению Риккати

$$-r'(a)=b(a)-r(a)^2, \quad r(T)=0. \tag{20.7}$$

Это уже задача Коши.

В рассмотренном нами случае нетрудно использовать любой метод. В общем же случае, в вариационных задачах большой размерности, существенно, что динамическое программирование приводит к решению задачи Коши для системы обыкновенных дифференциальных уравнений, которое можно легко реализовать на вычислительных машинах.

21. Инвариантное погружение

С другой стороны, мы можем рассматривать недостающее начальное условие в (20.2), т. е. $u'(a)$, как функцию параметра a. Оказывается, эта функция удовлетворяет обыкновенному дифференциальному уравнению типа Риккати. Такой подход имеет смысл использовать в более общем случае, когда рассматриваемое дифференциальное уравнение не является уравнением Эйлера вариационной задачи [1]).

22. Сочетание динамического программирования и квазилинеаризации

Используя метод динамического программирования, можно исследовать общую задачу минимизации функционала

$$J(u)=\int_a^T g(u', u, t)\,dt, \tag{22.1}$$

определенного на всех функциях, удовлетворяющих условию $u(a)=c$. Полагая

$$\min_u J(u)=f(a, c), \tag{22.2}$$

получаем нелинейное дифференциальное уравнение в частных производных

$$-\frac{\partial f}{\partial a}=\min_v\left[g(v, c, a)+v\frac{\partial f}{\partial c}\right] \tag{22.3}$$

с начальным условием $f(T, c)=0$. Численное решение этого уравнения не вызывает затруднений. Такой подход имеет определенные преимущества перед квазилинеаризацией и многими другими методами, поскольку длина отрезка $[a, T]$ играет незначительную роль, а наличие ограничений по существу упрощает численное решение.

Однако в случае больших размерностей численное решение наталкивается на затруднения, связанные с ограниченностью оперативной памяти. Чтобы обойти это препятствие, мы можем объединить локальную квазилинеаризацию с глобальным динамическим программированием. Этот вопрос мы обсудим в гл. VII.

---

[1]) Отметим, что линейные краевые задачи для дифференциальных уравнений можно свести к решению задач Коши для уравнений типа Риккати широко известным методом прогонки. — *Прим. ред.*

КОММЕНТАРИИ И БИБЛИОГРАФИЯ [1])

1. Мы совершенно не затронули очень трудную проблему существования и единственности решений двухточечных краевых задач. Читатель может ознакомиться с этим вопросом по книге

Коллатц Л., Численные методы решения дифференциальных уравнений, ИЛ, М., 1953.

4. О приложении матричных методов к линейным системам дифференциальных уравнений см.

Bellmann R., Introduction to matrix analysis, New York, 1960 (русский перевод готовится к печати в изд-ве «Наука»);
* Петровский И. Г., Лекции по теории обыкновенных дифференциальных уравнений, изд-во «Наука», М., 1964.
* Понтрягин Л. С., Обыкновенные дифференциальные уравнения, Физматгиз, М., 1961;
* Степанов В. В., Курс дифференциальных уравнений, Физматгиз, М., 1959.

5. О функциях Грина см.

Коддингтон Э. А., Левинсон Н., Теория обыкновенных дифференциальных уравнений, ИЛ, М., 1958.

6. Обсуждение роли выпуклости в анализе проводится в книге

Беккенбах Э., Беллман Р., Неравенства, изд-во «Мир», М., 1965,

где содержится много дополнительных ссылок.

7—10. Мы следуем работе

Kalaba R., On nonlinear differential equation, the maximum operation, and monotone convergence, *J. Math. Mech.*, **8** (1959), 519—574.

16. См.

Bellman R., Kalaba R., Dynamic programming, invariant imbedding and quasilinearization: comparisons and interconnections, в сборнике «Computing methods in optimization problems» под редакцией Balakrishnan A., Neustadt L., New York, 1964, стр. 135—145.

17. Здесь мы следуем работе

Bellman R., Quasilinearization and upper and lower bounds for variational problems, *Quart. Appl. Math.*, **19** (1962), 349—350.

Дальнейшие результаты можно найти в статьях

Hanson M. A., Bounds for functionally convex optimal control problems, *J. Math. Anal. Appl.*, **8** (1964), 84—89;
Bellman R., Functional equations and successive approximations in linear and nonlinear programming, *Naval Res. Logist Quart.*, **7** (1960), 63—83;

---

[1]) Работы, отмеченные звездочкой, добавлены при переводе. — *Прим. перев.*

Bellman R., An approximate procedure in control theory based on quasilinearization, *Bull. Univ. Jassi*, **10** (14), (1964), fasc. 3—4;
Bellman R., An application of dynamic programming to location-allocation problems, *SIAM Rev.*, **7** (1965), 126—128.

Мы используем идею двойственности, в каком-то смысле очень близкую идее Фридрихса; о приложении см.

Ikebe T., Kato T., Application of variational method to the Thomas-Fermi equation, *J. Phys. Soc. Japan*, **12** (1957), 201—204.

19. Детальное обсуждение и некоторые дальнейшие приложения содержатся в работах

Беллман Р., Гликсберг И., Гросс О., Некоторые вопросы математической теории процессов управления, ИЛ, М., 1962;
Bellman R., Some nonclassical problems in the calculus of variations, *Proc. Amer. Math. Soc.*, **7** (1956), 87—94.

20. О различных применениях этого подхода см.

Беллман Р., Дрейфус С., Прикладные задачи динамического программирования, изд-во «Наука», М., 1965;
Беллман Р., Процессы регулирования с адаптацией, изд-во «Наука», М., 1964.

21. Об инвариантном погружении см.

Bellman R., Kalaba R., Wing G. M., Invariant imbedding and mathematical physics — I, Particle processes, *J. Math. Phys.*, **1** (1960), 280—308;
Bellman R., Kalaba R., Prestrud M., Invariant imbeddnig and radiative transfer in slabs of finite thickness, New York, 1963;
Bellman R., Kagiwada H., Kalaba R., Prestrud M., Invariant imbedding and time-dependent transport processes, New York, 1964.
Калаба Р., Инвариантное погружение и анализ процессов, сб. «Общая теория систем» под редакцией Месаровича М., изд-во «Мир», М., 1966.

Глава III

Монотонное поведение и дифференциальные неравенства

1. Монотонность

Сосредоточим наше внимание на свойствах монотонности последовательности приближений, полученной с помощью квазилинеаризации. Исследование этих свойств важно по нескольким причинам. Во-первых, эти результаты важны с аналитической точки зрения, так как они дают новые методы доказательства ограниченности и сходимости приближений и попутно, как мы уже отмечали при изучении уравнения Риккати, позволяют найти максимальный интервал сходимости. Во-вторых, монотонная сходимость весьма полезна с вычислительной точки зрения, так как она позволяет получить границы решения и дает возможность контролировать численные результаты.

Изучение монотонного поведения приближений для случая нелинейных дифференциальных уравнений второго порядка эквивалентно изучению дифференциальных неравенств вида

$$u'' + p(t)\,u' + q(t)\,u \geqslant 0. \tag{1.1}$$

Исследование условий, при которых из этих неравенств вытекает $u \geqslant 0$ или $u \leqslant 0$, приводит нас в классические области анализа, из которых одни настраивают нас оптимистически, а другие вызывают лишь уныние. Мы воспротивимся искушению совершить продолжительный экскурс в какую-либо из этих областей и сосредоточим наше внимание на результатах, наиболее необходимых на данном этапе. В конце главы читатель найдет много ссылок на дальнейшие результаты.

Применяя квазилинеаризацию к уравнению Риккати, мы неоднократно использовали факт существования простого точного решения линейного уравнения первого порядка

$$u' = a(t)\,u + b(t). \tag{1.2}$$

Это решение, выраженное через квадратуры и экспоненциальные функции, позволило нам получить всю информацию, какую мы хотели, о дифференциальном неравенстве

$$u' - a(t)\,u - b(t) \geqslant 0. \tag{1.3}$$

При изучении (1.1) мы не располагаем такой роскошью и, следовательно, должны проявить некоторую изобретательность.

В данной главе мы представим целый набор полезных для нас приемов и хитроумных уловок, почерпнутых из вариационного исчисления и теории дифференциальных уравнений в частных производных.

Для начала отметим, что достаточно, вообще говоря, изучить неравенство

$$u'' + q(t)u \geqslant 0, \tag{1.4}$$

так как сохраняющая знак замена переменной $u = e^{-\int p\,dt/2} v$ превращает (1.1) в неравенство

$$v'' + \left(q(t) - \frac{p'}{2} - \frac{p^2}{4}\right) v \geqslant 0. \tag{1.5}$$

2. Элементарный подход

Прежде чем прибегнуть к некоторым высокоэффективным методам, дадим прямой элементарный метод изучения неравенства

$$u'' - a(t)u \geqslant 0, \quad 0 \leqslant t \leqslant b. \tag{2.1}$$

Этот метод может применяться при исследовании обыкновенных нелинейных дифференциальных уравнений, эллиптических и параболических дифференциальных уравнений в частных производных, а также других типов функциональных уравнений.

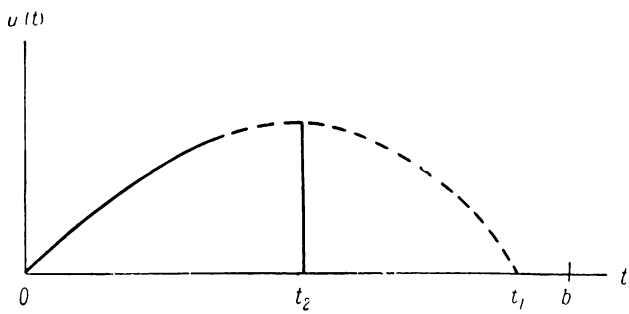


Рис. 3.1.

Пусть $a(t) > 0$ на $[0, b]$ и $u(0) \geqslant 0$, $u'(0) > 0$. Мы хотим показать, что в этом случае $u(t) \geqslant 0$ при $0 < t \leqslant b$. Предположим, что этот результат несправедлив. Тогда $u(t)$ обращается где-то в нуль, например в точке $t = t_1$, как показано на рис. 3.1. Следовательно, существует промежуточная точка t_2, в которой $u(t)$ имеет относительный максимум. Однако в этой точке

$$u''(t) \geqslant a(t)u(t) > 0, \tag{2.2}$$

а это противоречит утверждению, что u имеет локальный максимум при $t=t_2$.

Итак, мы доказали, что из неравенств $u(0)\geqslant 0$, $u'(0)>0$ следует $u(t)>0$ при $0<t\leqslant b$.

3. Связи между решениями и коэффициентами уравнения

Чтобы положить начало более глубокому анализу дифференциального неравенства (2.1), рассмотрим некоторые связи между решениями и коэффициентами уравнения, аналогичные связям между корнями и коэффициентами полиномиальных уравнений.

Пусть u_1 и u_2 — два линейно независимых решения уравнения

$$u''+p(t)u'+q(t)u=0. \tag{3.1}$$

Рассмотрим детерминантное уравнение

$$\begin{vmatrix} u_1 & u_1' & u_1'' \\ u_2 & u_2' & u_2'' \\ u & u' & u'' \end{vmatrix}=0. \tag{3.2}$$

Так как это линейное дифференциальное уравнение второго порядка, имеющее два линейно независимых решения $u=u_1$ и $u=u_2$, то оно должно быть эквивалентно исходному уравнению.

Вводя обозначение

$$W(u_1, u_2)=\begin{vmatrix} u_1 & u_1' \\ u_2 & u_2' \end{vmatrix} \tag{3.3}$$

(вронскиан от u_1 и u_2), можно записать (3.1) в виде

$$u''+u'\frac{\begin{vmatrix} u_1'' & u_1 \\ u_2'' & u_2 \end{vmatrix}}{W(t)}+u\frac{\begin{vmatrix} u_1' & u_1'' \\ u_2' & u_2'' \end{vmatrix}}{W(t)}=0 \tag{3.4}$$

(мы пишем, как и раньше, $W(t)\equiv W(u_1, u_2)$). Сравнивая (3.4) и (3.1), получаем искомые выражения для $p(t)$ и $q(t)$ через u_1 и u_2.

Чтобы убедиться, что наличие $W(t)$ в знаменателе не может привести к особенности, выведем результат Абеля

$$W(u_1, u_2)=W(u_1(0), u_2(0))\,e^{-\int_0^t p(s)\,ds}. \tag{3.5}$$

Дифференцируя (3.3), имеем

$$\frac{dW}{dt} = \frac{d}{dt}(u_1u_2' - u_2u_1') = u_1u_2'' - u_2u_1'' =$$
$$= (p(t)u_1' + q(t)u_1)u_2 - u_1(p(t)u_2' + q(t)u_2) =$$
$$= -p(t)(u_1u_2' - u_1'u_2) = -p(t)\,W, \quad (3.6)$$

откуда сразу следует (3.5). Из соотношения (3.5) получаем, что из линейной независимости решений u_1 и u_2 в одной точке, скажем при $t=0$, следует их линейная независимость во всех точках. Если в качестве u_1 и u_2 выбрать фундаментальную систему решений, удовлетворяющую начальным условиям $u_1(0)=1$, $u_1'(0)=0$, $u_2(0)=0$, $u_2'(0)=1$, то мы получим $W(u_1(0), u_2(0))=1$. Следовательно, определитель $W(t)$ положителен при $t\geqslant 0$ и отличен от нуля, если только функция $\int_0^t p(s)ds$ конечна при $t\geqslant 0$. Мы уже использовали этот результат в разд. 3 гл. II.

4. Факторизация линейного дифференциального оператора второго порядка

Проделаем теперь некоторые простые преобразования, которые позволят нам произвести весьма полезное разбиение уравнения

$$u'' + p(t)u' + q(t)u = 0, \quad (4.1)$$

или, что то же, оператора $D^2+p(t)D+q(t)$, где $D=d/dt$. Будем исходить из соотношений

$$\frac{d}{dt}\left(\frac{u}{u_1}\right) = \frac{\begin{vmatrix} u_1 & u_1' \\ u & u' \end{vmatrix}}{u_1^2} \quad (4.2)$$

и

$$\frac{d}{dt}\left(\frac{\begin{vmatrix} u_1 & u_1' \\ u & u' \end{vmatrix}}{\begin{vmatrix} u_1 & u_1' \\ u_2 & u_2' \end{vmatrix}}\right) = \frac{\begin{vmatrix} u_1 & u_1' \\ u_2 & u_2' \end{vmatrix}\begin{vmatrix} u_1 & u_1'' \\ u & u'' \end{vmatrix} - \begin{vmatrix} u_1 & u_1' \\ u & u' \end{vmatrix}\begin{vmatrix} u_1 & u_1'' \\ u_2 & u_2'' \end{vmatrix}}{\begin{vmatrix} u_1 & u_1' \\ u_2 & u_2' \end{vmatrix}^2}. \quad (4.3)$$

Мы предполагаем, что u_1 и u_2 таковы, что $W(u_1, u_2)\neq 0$. Используем теперь некоторые простые результаты теории определителей. Пусть

$$A = (a_{ij}) \quad (4.4)$$

— матрица 3×3, имеющая обратную матрицу

$$A^{-1}=(A_{ij}/|A|)', \tag{4.5}$$

где A_{ij} — алгебраическое дополнение элемента a_{ij}, а $|A|$ обозначает определитель матрицы A. Так как $(A^{-1})^{-1}=A$, то

$$a_{11}=\frac{\begin{vmatrix} A_{22}/|A| & A_{32}/|A| \\ A_{23}/|A| & A_{33}/|A| \end{vmatrix}}{|A^{-1}|}, \tag{4.6}$$

или

$$a_{11}|A|=\begin{vmatrix} A_{22} & A_{32} \\ A_{23} & A_{33} \end{vmatrix}. \tag{4.7}$$

Рассмотрим теперь определитель

$$W(u_1, u_2, u)=\begin{vmatrix} u_1 & u_1' & u_1'' \\ u_2 & u_2' & u_2'' \\ u & u' & u'' \end{vmatrix}. \tag{4.8}$$

Обозначая через A_{ij} алгебраическое дополнение элемента с номером ij в этом определителе и используя (4.7), имеем

$$A_{33}A_{22}-A_{23}A_{32}=u_1 W(u_1, u_2, u). \tag{4.9}$$

Отождествляя эти алгебраические дополнения с членами, стоящими в правой части соотношения (4.3), получаем

$$\frac{d}{dt}\left(\frac{\begin{vmatrix} u_1 & u_1' \\ u & u' \end{vmatrix}}{\begin{vmatrix} u_1 & u_1' \\ u_2 & u_2' \end{vmatrix}}\right)=\frac{u_1 W(u_1, u_2, u)}{W(u_1, u_2)^2}. \tag{4.10}$$

Следовательно [1]),

$$\frac{W(u_1, u_2, u)}{W(u_1, u_2)}=\frac{W(u_1, u_2)}{u_1}\frac{d}{dt}\left(\frac{W(u_1, u)}{W(u_1, u_2)}\right)=$$

$$=\frac{W(u_1, u_2)}{u_1}\frac{d}{dt}\left(\frac{u_1^2}{W(u_1, u_2)}\frac{d}{dt}\left(\frac{u}{u_1}\right)\right) \tag{4.11}$$

[1]) При выводе соотношения (4.11) использовано тождество (4.2). Равенство (4.12) следует из того, что уравнения (3.1) и (3.2) эквивалентны (см. разд. 3), причем левая часть уравнения (3.2) с учетом обозначений (3.3) и (4.8) равна левой части соотношения (4.11). — *Прим. ред.*

и окончательно имеем искомое разбиение

$$u''+p(t)u'+q(t)u=\frac{W(u_1,u_2)}{u_1}\frac{d}{dt}\left(\frac{u_1^2}{W(u_1,u_2)}\frac{d}{dt}\left(\frac{u}{u_1}\right)\right). \quad (4.12)$$

В следующей главе мы приведем аналогичные результаты для линейного дифференциального уравнения N-го порядка.

5. Свойство положительности

С помощью последнего результата предыдущего раздела (4.12) можно легко доказать следующее утверждение. Для того чтобы из неравенства

$$u''+p(t)u'+q(t)u\geqslant 0, \quad u(0)=c_1, \quad u'(0)=c_2, \quad (5.1)$$

при $0\leqslant t\leqslant b$ следовало неравенство $u\geqslant v$, где v удовлетворяет дифференциальному уравнению

$$v''+p(t)v'+q(t)v=0, \quad v(0)=c_1, \quad v'(0)=c_2, \quad (5.2)$$

достаточно, чтобы существовало положительное на $[0, b]$ решение однородного уравнения

$$w''+p(t)w'+q(t)w=0.$$

Обозначим это решение через $u_1(t)$, и пусть $u_2(t)$ — другое решение этого уравнения, причем $u_1(t)$ и $u_2(t)$ линейно независимы. Тогда определитель $W(u_1, u_2)$ не равен нулю и можно выбрать u_2 так, чтобы сделать этот определитель положительным. Как мы знаем, этого можно добиться выбором только $u_2(0)$ и $u_2'(0)$. Из (5.1) и (5.2) заключаем, что

$$(u-v)''+p(t)(u-v)'+q(t)(u-v)=r(t), \quad (5.3)$$

где $r(t)\geqslant 0$ при $0\leqslant t\leqslant b$. С помощью (4.12) можно записать (5.3) в виде

$$\frac{d}{dt}\left(\frac{u_1^2}{W(u_1,u_2)}\frac{d}{dt}\left(\frac{u-v}{u_1}\right)\right)=\frac{u_1 r(t)}{W(u_1,u_2)}, \quad (5.4)$$

где u_1 и u_2 определены выше. Интегрируя (5.4) по отрезку $[0, t]$, получаем

$$\frac{u_1^2}{W(u_1,u_2)}\frac{d}{dt}\left(\frac{u-v}{u_1}\right)=\int_0^t\frac{u_1 r(s)\,ds}{W(u_1,u_2)}\geqslant 0, \quad (5.5)$$

так как по предположению u_1 и r неотрицательны и $W>0$. Константа интегрирования в (5.5) равна нулю, поскольку $u(0)=$

$=v(0)$, $u'(0)=v'(0)$, и поэтому $\frac{d}{dt}\left(\frac{u-v}{u_1}\right)=0$ при $t=0$. Интегрируя еще раз, получаем

$$\frac{u-v}{u_1}=\int_0^t\left[\frac{W(u_1, u_2)}{u_1^2}\int_0^s\frac{u_1 r(s_1)\,ds_1}{W(u_1, u_2)}\right]ds. \tag{5.6}$$

Здесь снова константа интегрирования равна нулю, так как $u(0)=v(0)$. Отсюда получаем результат $u\geqslant v$, к которому мы стремились. Кроме того, имеем оценку снизу для величины $u-v$, выраженную через $r=L(u)-L(v)$, где L — линейный оператор второго порядка. Наконец, отметим, что существование $u_1(t)$ связано с условием на собственные значения.

6. Сопутствующее параболическое дифференциальное уравнение в частных производных

Сейчас мы хотим исследовать вопрос о положительности, или точнее неотрицательности, решения уравнения

$$u''+q(x)u+f(x)=0, \quad 0<x<b, \tag{6.1}$$

с граничными условиями $u(0)=u(b)=0$, изучив для этого поведение решения связанного с ним уравнения теплопроводности

$$u_t=u_{xx}+q(x)u+f(x) \tag{6.2}$$

с начальным условием

$$u(x, 0)=h(x), \quad 0\leqslant x\leqslant b, \tag{6.3}$$

и с граничными условиями

$$u(0, t)=u(b, t)=0, \quad t>0. \tag{6.4}$$

Для удобства мы изменили обозначение переменных, чтобы сохранить за x привычное обозначение пространственной переменной, а за t — временно́й переменной.

Обоснованием введения уравнения (6.2) служит следующий факт. При соответствующих условиях, которые мы исследуем ниже, предельное значение $u(x, t)$ при $t\to\infty$ является решением уравнения (6.1). Следовательно, если мы сможем доказать неотрицательность решения уравнения (6.2) для $t\geqslant 0$, то мы докажем неотрицательность решения уравнения (6.1). Мы пользуемся здесь методом *погружения*.

Предположим, что мы доказали необходимые теоремы существования и единственности и сейчас нас интересует только положительность решения.

Для начала рассмотрим уравнение

$$u_t = u_{xx} + f(x), \tag{6.5}$$

которое порождается моделью диффузии или случайного блуждания. Соответствующая разностная схема имеет вид

$$u\left(x,\ t+\frac{\Delta^2}{2}\right) = \frac{u(x+\Delta,\ t)+u(x-\Delta,\ t)}{2} + f(x)\frac{\Delta^2}{2}, \tag{6.6}$$

где $x=0, \Delta, \ldots$ и $t=0,\ \Delta^2/2, \ldots$. Мы определяем функцию $u(x,\ t)\equiv u(x,\ t,\ \Delta)$ для всех $x,\ t\geqslant 0$, используя линейную интерполяцию между узлами сетки.

Непосредственно из рекуррентного соотношения следует, что $u(x,\ t)\geqslant 0$ для всех x и $t\geqslant 0$, если $f(x)\geqslant 0$ и $u(x,\ 0)\geqslant 0$. Тем самым показана неотрицательность решения уравнения (6.5), если, конечно, мы сможем доказать, что предел при $\Delta\to 0$ решения $u(x,\ t,\ \Delta)$ уравнения (6.6) является решением уравнения (6.5).

Для исследования более общего уравнения

$$u_t = u_{xx} + q(x,\ t)\,u + f(x), \tag{6.7}$$

когда $q(x,\ t)\geqslant 0$ при $x,\ t\geqslant 0$, напишем вместо (6.6) соотношение

$$u\left(x,\ t+\frac{\Delta^2}{2}\right) = \frac{u(x+\Delta,\ t)+u(x-\Delta,\ t)}{2} + \\ + q(x,\ t)\,u(x,\ t)\frac{\Delta^2}{2} + f(x)\frac{\Delta^2}{2}. \tag{6.8}$$

Здесь опять неотрицательность $u(x,\ t)$ устанавливается легко. Интересно отметить, что ограничение $q(x,\ t)\geqslant 0$ несущественно, так как можно всегда сделать замену переменной

$$u = e^{-kt}v, \tag{6.9}$$

которая приводит к уравнению

$$v_t = v_{xx} + (q(x,\ t) + k)\,v + f(x)\,e^{kt}. \tag{6.10}$$

Следовательно, если можно найти такую константу k, что $q(x,\ t)+k\geqslant 0$ для всех x и t, то доказательство проводится аналогично.

Предполагая, как и выше, что доказательство того факта, что предел решения разностного уравнения является решением дифференциального уравнения, не вызывает затруднений, мы, таким образом, получаем следующий результат.

Если

а) $u_t - u_{xx} - q(x, t)u \geqslant 0, \quad 0 < x < 1, \quad t > 0;$

б) $u(x, 0) \geqslant 0, \quad 0 \leqslant x \leqslant 1;$

в) $u(0, t) = u(1, t) = 0, \quad t > 0;$

г) $q(x, t) \geqslant -k(T) > -\infty, \quad 0 \leqslant x \leqslant 1, \quad 0 \leqslant t \leqslant T,$

для любого $T > 0$,

то $u(x, t) \geqslant 0$ при $0 \leqslant x \leqslant 1$, $t \geqslant 0$.

Отметим, что вопросам, связанным с доказательством сходимости дискретного процесса к непрерывному, посвящено большое число работ.

7. Еще о собственных значениях

Теперь возникает интересный вопрос, почему в случае дифференциального уравнения в частных производных решение положительно при довольно слабых требованиях, в то время как в случае обыкновенных дифференциальных уравнений решение положительно только при выполнении специальных условий. Как мы увидим, это соответствует истинному состоянию дел, а не является простым следствием нашего аналитического подхода.

Пусть уравнение

$$u'' + q(x)u + f(x) = 0, \quad u(0) = u(1) = 0, \tag{7.1}$$

имеет единственное решение $u = u_0(x)$, и пусть $u(x, t)$ — решение уравнения

$$\begin{gathered} u_t - u_{xx} - q(x)u - f(x) = 0, \\ u(0, t) = u(1, t) = 0, \quad u(x, 0) = h(x), \quad 0 \leqslant x \leqslant 1. \end{gathered} \tag{7.2}$$

Если положить

$$u(x, t) = u_0(x) + v(x, t), \tag{7.3}$$

то видно, что v удовлетворяет однородному уравнению

$$\begin{gathered} v_t - v_{xx} - q(x)v = 0, \\ v(0, t) = v(1, t) = 0, \quad v(x, 0) = h(x) - u_0(x). \end{gathered} \tag{7.4}$$

Для того чтобы $u(x, t)$ сходилось к $u_0(x)$ при $t \to \infty$, необходимо, чтобы $v(x, t) \to 0$ при $t \to \infty$. Используя метод разделения переменных для решения краевой задачи (7.4), получаем [1])

$$v(x, t) = \sum_{k=1}^{\infty} c_k e^{\lambda_k t} v_k(x). \tag{7.5}$$

[1]) Произвольные постоянные c_k в разложении (7.5) определяются при помощи начального условия (7.4). — *Прим. ред.*

Пары λ_k и $v_k(x)$ определяются из задачи Штурма — Лиувилля

$$\lambda_k v_k - v_k'' - q(x) v_k = 0, \quad v_k(0) = v_k(1) = 0, \tag{7.6}$$

где $v_k(x)$ нормировано согласно условию $\int_0^1 v_k^2(x)\,dx = 1$. Если все собственные значения λ_k задачи (7.6) отрицательны, то $v(x, t) \to 0$ при $t \to \infty$. Только в этом случае $u(x, t) \to u(x)$ при $t \to \infty$ для *всех* вынуждающих функций $f(x)$, и, следовательно, только при наличии такого ограничения на собственные значения можно утверждать, что из неотрицательности $u(x, t)$ следует неотрицательность $u(x)$.

8. Вариационный подход

Изучение дифференциального неравенства

$$u'' + q(t) u \leqslant 0, \quad 0 \leqslant t \leqslant b, \quad u(0) = u(b) = 0, \tag{8.1}$$

эквивалентно изучению дифференциального уравнения

$$u'' + q(t) u + f(t) = 0, \quad 0 \leqslant t \leqslant b, \quad u(0) = u(b) = 0, \tag{8.2}$$

для произвольной неотрицательной функции $f(t)$ [1]).

Сейчас мы опишем метод (целиком отличный от всех трех изложенных выше) исследования уравнения (8.2), основанный на том факте, что оно является уравнением Эйлера, связанным с задачей минимизации функционала

$$J(u) = \int_0^b \left[u'^2 - q(t) u^2 - 2 f(t) u\right] dt, \tag{8.3}$$

заданного на функциях $u(t)$, которые удовлетворяют условиям $u(0) = u(b) = 0$ и для которых интеграл (8.3) существует.

Следовательно, для того чтобы продемонстрировать неотрицательность решения уравнения (8.2) при $f(t) \geqslant 0$, достаточно показать, что функция, минимизирующая $J(u)$, неотрицательна, если $f(t) \geqslant 0$.

Для начала потребуем, чтобы

$$q(t) \leqslant \frac{\pi^2}{b^2} - \varepsilon, \quad \varepsilon > 0, \quad 0 \leqslant t \leqslant b \tag{8.4}$$

(условие на собственные значения, см. разд. 9). Это условие гарантирует наличие у $J(u)$ конечной нижней грани. Стандартные рассуждения, основанные на выпуклости, показывают, что ниж-

[1]) Мы снова вернулись к обозначению независимой переменной через t.

няя грань достигается на единственной функции. Ссылки на работы, где содержатся эти результаты, читатель найдет в конце данной главы.

Чтобы показать неотрицательность минимизирующей функции, проведем доказательство методом от противного. Предположим, что $u(t) < 0$ на отрезке $[a_1, b_1]$, лежащем внутри отрезка $[0, b]$, как показано на рис. 3.2. Затем рассмотрим новую функцию $w(t)$, равную $-u(t)$ на отрезке $[a_1, b_1]$ и равную $u(t)$ вне

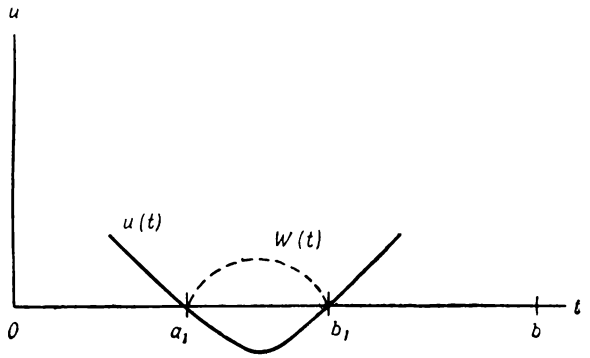


Рис. 3.2.

его. Фукция $w'(t)$, возможно, будет разрывна, но все равно квадратично интегрируема. Такое изменение не влияет на квадратичные члены в (8.3), но уменьшает вклад члена, содержащего $f(t)$. Это противоречит тому, что $u(t)$ — минимизирующая функция.

Заметим, что из этого доказательства видно, что нет необходимости искать решение в классе функций, принимающих отрицательные значения, даже если мы не доказали предварительно единственность решения.

9. Обсуждение

Уравнение

$$u'' + ku + \sin\frac{\pi t}{b} = 0, \quad u(0) = u(b) = 0, \tag{9.1}$$

с решением

$$u = \frac{-\sin\frac{\pi t}{b}}{k - \frac{\pi^2}{b^2}} \tag{9.2}$$

показывает, что оценку (8.4) нельзя улучшить, если имеется в виду отыскание *равномерной* границы для $q(t)$, не зависящей от $f(t)$. Если $k>\pi^2/b^2$, то знак u противоположен знаку $\sin \pi t/b$.

Величина π^2/b^2 является наименьшим собственным значением для задачи Штурма — Лиувилля

$$u''+\lambda u=0, \quad u(0)=u(b)=0; \tag{9.3}$$

этим еще раз демонстрируется связь с собственными значениями.

Изложенное выше доказательство можно использовать для получения более общего результата, так называемого свойства осцилляции, состоящего в том, что u не может иметь число перемен знака больше, чем $f(t)$.

Решение уравнения (8.2) можно записать в виде

$$u=\int_0^b k(t, s) f(s)\, ds, \tag{9.4}$$

где k — функция Грина. Приведенные ранее доказательства устанавливают неотрицательность функции Грина и фактически ее свойство осцилляции. С более обширными результатами, касающимися свойства осцилляции функции Грина, читатель может познакомиться в литературе, указанной в конце главы.

Отметим, что доказательство в разд. 8 можно распространить на общую вариационную задачу с функционалом

$$J(u)=\int_0^b [g(u, u')-2f(t)u]\, dt, \tag{9.5}$$

где g — четная функция от u и u'.

Наконец, еще раз укажем, что мы постоянно воздерживались от рассмотрения уравнения более общего вида

$$u''+p(t)u'+q(t)u=0, \tag{9.6}$$

поскольку замена переменной

$$u=e^{-\frac{1}{2}\int_0^t p(s)\, ds} v \tag{9.7}$$

превращает (9.6) в уравнение

$$v''+\left(q(t)-\frac{p'(t)}{2}-\frac{p^2(t)}{4}\right)v=0. \tag{9.8}$$

10. Улучшение оценок сходимости

Фактически в теории дифференциальных (и вообще функциональных) уравнений наиболее желателен метод последовательных приближений, который обеспечивал бы быструю сходимость к решению всегда, когда оно существует, причем во всей области существования. Покажем, что установленный выше факт монотонности позволяет в некоторых простых случаях доказать это свойство для метода квазилинеаризации. Полное исследование увело бы нас слишком далеко в общую теорию дифференциальных неравенств.

Рассмотрим, например, изученное в разд. 12 гл. II уравнение

$$u'' = e^u, \quad u(0) = u(b) = 0. \tag{10.1}$$

Последовательность приближений порождается соотношением

$$u''_{n+1} = e^{u_n} + (u_{n+1} - u_n) e^{u_n}, \tag{10.2}$$

где для удобства выбрано $u_0 = 0$. Так как $e^{u_n} > 0$ для всех n, то из результатов разд. 2 следует, что имеет место монотонность

$$u_1 \geqslant u_2 \geqslant \ldots \geqslant u_n \geqslant \ldots \geqslant u \tag{10.3}$$

Пусть известно, что существует единственное решение. Тогда мы делаем вывод, что последовательность $\{u_n\}$, как монотонно ограниченная последовательность, обязательно сходится. Более того, из предыдущего известно, что она сходится к u. Так как функция u непрерывна и сходимость монотонна, то сходимость должна быть *равномерной* (теорема Дини). Этот факт мы также обсуждали ранее.

Таким образом, положительность оператора дает новый метод доказательства сходимости, а кроме того, и улучшенные оценки области сходимости.

Вообще, рассматривая уравнение

$$u'' = g(u, x), \quad u(0) = u(b) = 0, \tag{10.4}$$

и связанные с ним линейные уравнения

$$u''_{n+1} = g(u_n, x) + (u_{n+1} - u_n) g_u(u_n, x), \tag{10.5}$$

порождающие последовательность приближений, мы видим, что монотонность зависит от справедливости неравенства

$$-g_u(u_n, x) \leqslant \frac{\pi^2}{b^2} - \varepsilon. \tag{10.6}$$

Если функция g_u всегда положительна, как это имело место выше, или если она монотонна, выпукла или вогнута, так что исследование $-g_u(u_1, x)$ или $-g_u(u, x)$ помогает сделать определенные выводы, то можно установить более общие результаты.

Мы хотели просто показать некоторые возможности этого подхода, не вдаваясь слишком глубоко в аналитические аспекты проблемы.

11. Обсуждение

Мы использовали три различных подхода к изучению поведения функции Грина для линейного дифференциального уравнения второго порядка

$$u'' + p(t)u' + q(t)u + f(t) = 0, \quad u(0) = u(b) = 0. \tag{11.1}$$

Первый подход, рассмотренный в разд. 2—5, был прямым. Второй и третий подходы были косвенными, основанными на связи уравнения (11.1) с другими процессами и соответствующими им уравнениями. Подробно исследуя эти процессы, мы доказывали положительность решения и делали другие выводы.

Отметим также, что мы использовали методы погружения. Исходная задача была погружена в более широкое множество задач, которые фактически решаются легче. Эта идея является фундаментальной в динамическом программировании и инвариантном погружении — методах, которые мы обсудили кратко в гл. II. Мы вновь вернемся к динамическому программированию в гл. VII.

КОММЕНТАРИИ И БИБЛИОГРАФИЯ

1. Систематическое изучение дифференциальных неравенств восходит к работе С. А. Чаплыгина. См., например,

Чаплыгин С. А., Новый метод приближенного интегрирования дифференциальных уравнений, Собр. соч., Гостехиздат, 1948, т. 1, стр. 347—419.

Эти исследования приводят непосредственно к изучению обобщенной выпуклости, которой занимались Беккенбах, Бинг, Валирон, Бонсалл, Пексото, Моцкин, Торнхейм, Куртис, Рейд, Хартман, Редхеффер и др. Ссылки на литературу можно найти на стр. 219 книги

Беккенбах Э., Беллман Р., Неравенства, изд-во «Мир», М., 1965.

См. также

Красносельский М. А., Положительные решения операторных уравнений, Гостехиздат, М., 1961;
Коллатц Л., Численные методы решения дифференциальных уравнений, ИЛ, М., 1953.

2. Дальнейшие примеры, в которых могут использоваться доказательства такого типа, можно найти в гл. 6 и 7 книги

Беллман Р., Теория устойчивости решений дифференциальных уравнений, ИЛ, М., 1954.

О применении подобных рассуждений к гармоническим функциям, т. е. к решениям уравнения $u_{xx}+u_{yy}=0$, см.

Титчмарш Е., Теория функций, ГИТТЛ, М.—Л., 1951.

3. Соответствующие результаты для линейных дифференциальных уравнений n-го порядка получены в работах

Polya G., On the mean-value theorem corresponding to a given linear homogeneous differential equation, *Trans. Amer. Math. Soc.*, **24** (1922), 312—324;
Schlesinger L., Handbuch der Theorie der Linearen Differentialgleichungen, т. 1, 1895, стр. 52.

5. Этот результат неоднократно получали многие авторы. Оказывается, он восходит к С. А. Чаплыгину. См. ссылку в разд. 1, а также

Wilkins J. E., The converse of a theorem of Tchaplygin on differential equations, *Bull. Amer. Math. Soc.*, **53** (1947), 126—129.

6. Что касается обсуждения условий, при которых дискретная схема сходится к непрерывной, см.

Форсайт Дж., Вазов В., Разностные методы решения дифференциальных уравнений в частных производных, ИЛ, М., 1963.

Имеется много обобщений этих методов и связанных с ними приложений. Например, можно заменить уравнение (6.6) уравнением

$$u(x,\ t+\Delta^2)=\sum_{i=1}^{N} a_i\,[u(x+w_i\Delta,\ t)+u(x-w_i\Delta,\ t)]+\ \ldots$$

и путем здравого выбора a_i и w_i получить гораздо бóльшую степень точности. Кроме того, можно применить ту же самую идею к нелинейным уравнениям в частных производных и таким образом доказать неотрицательность решения. Например, $u_t=uu_x$ можно аппроксимировать уравнением

$$u(x,\ t+\Delta)=u(x+u(x,\ t)\Delta,\ t).$$

См.

Bellman R., Cherry I., Wing G. M., A note on the numerical integration of a class of nonlinear hyperbolic equations, *Quart. Appl. Math.*, **16** (1958), 181—183;
Bellman R., Some questions concerning difference approximations to partial differential equations, *Boll. d'Unione Mate.*, **17** (1962), 188—190;
Bellman R., Kalaba R., Kotkin B., On a new approach to the computational solution of partial differential equations, *Proc. Nat. Acad. Sci. USA*, **48** (1962), 1325—1327;
Motzkin T., Wasow W., On the approximation of linear elliptic partial differential equations by difference equations with positive coefficients, *J. Math. Phys.*, **31** (1953), 253—259.

Мимоходом отметим применение квазилинеаризации к изучению решения уравнения теплопроводности, см.

Bellman R., Some properties of summation kernels, *Duke Math. J.*, **15** (1948), 1013—1019.

Здесь мы используем результат

$$\max_{a \leqslant x \leqslant b} |u(x)| = \lim_{n \to \infty} \left(\int_a^b u(x)^{2n} \, dx \right)^{1/2n}$$

8. Этот подход впервые был изложен в работе

Bellman R., On the nonnegativity of Green's functions, *Boll. d'Unione Mate.*, **12** (1957), 411—413.

О доказательстве существования минимума см.

Беллман Р., Гликсберг И., Гросс О., Некоторые вопросы математической теории процессов управления, ИЛ, М., 1962.

9. См.

Bellman R., On variation-diminishing properties of Green's functions, *Boll. d'Unione Mate.*, **16** (1961), 164—166.

Что касается детального обсуждения преобразований, уменьшающих осцилляцию, и понятия полной положительности, см.

Karlin S., Total positivity and convexity preserving transformations, *Proc. Symp. Pure Math.*, **7** (1963), 329—347.

Другие ссылки приведены в книге, упомянутой в разд. 1.

10. Относительно обобщения теоремы Дини на L^p-пространство и даже на более общие пространства см.

Vala K., On compact sets of compact operators, *Ann. Acad. Sci. Fenn., Ser. A*, **351** (1946), 3—8;
Mlak W., Note on abstract differential inequalities and Chaplighin method, *Ann. Polon Math.*, **10** (1961), 253 ff.;
Mlak W., Note on maximum solutions of differential equations, *Contributions to differential equations*, **1** (1963), 461—465.

Вот некоторые представляющие особый интерес статьи о дифференциальных неравенствах:

Collatz L., Aufgaben monotoner Art, *Arch. Math.*, **3** (1952), 366—376;
Collatz L., Applications of the theory of monotonic operators to boundary value problems, *Boundary problems in differential equations*, University of Wisconsin Press, Madison, Wisconsin, 1960;
Redheffer R. M., Die Collatzsche Monotonie bei Anfangswertproblemen, *Arch. Rational Mech. Anal.*, **14** (1963), 196—212;
Авербух А. И., Связь теоремы С. А. Чаплыгина с теорией оптимальных процессов, *Автоматика и телемеханика*, **22**, № 10 (1961), 1309—1313;
Jones G. S., Fundamental inequalities for discrete and discontinuous functional equations, *J. Soc. Indust Appl. Math.*, **12** (1964), 43—57.

Глава IV

Системы дифференциальных уравнений, память и дифференциальная аппроксимация

1. Введение

В этой главе мы намереваемся обсудить приложение наших методов к исследованию двухточечных и многоточечных краевых задач для дифференциальных уравнений высшего порядка и систем дифференциальных уравнений. Как и раньше, вначале мы изложим метод квазилинеаризации, обеспечивающий квадратичную сходимость, а затем рассмотрим вопросы монотонности.

Вычислительные аспекты, особенно проблема экономии оперативной памяти, потребуют от нас изыскания нового подхода к рассмотрению последовательных приближений. Это в свою очередь приведет нас к новой методике получения численного решения дифференциально-разностных уравнений и более общих функционально-дифференциальных уравнений типа

$$u'(t) = g(t,\ u(t),\ u(h(t))). \tag{1.1}$$

Мы используем также квазилинеаризацию в связи с дифференциальной аппроксимацией, что позволит нам свести довольно сложные функциональные уравнения к обыкновенным дифференциальным уравнениям. Будут приведены примеры, иллюстрирующие эффективность такой процедуры.

Нелинейное дифференциальное уравнение

$$u^{(N)} = g(u,\ u',\ \ldots,\ u^{(N-1)}) \tag{1.2}$$

можно непосредственно исследовать с помощью квазилинеаризации. Однако более удобно как в аналитических, так и в вычислительных целях заменить это уравнение системой уравнений первого порядка путем введения новых переменных $u_1, u_2, \ldots, u_N$, где $u_1 = u$, $u_2 = u', \ldots, u_N = u^{(N-1)}$. Тогда (1.2) примет вид

$$\begin{aligned} u_1' &= u_2, \\ u_2' &= u_3, \\ &\cdot \\ &\cdot \\ &\cdot \\ u_N' &= g(u_1,\ u_2,\ \ldots,\ u_N). \end{aligned} \tag{1.3}$$

Как мы увидим ниже, при исследовании вопросов монотонности в некоторых случаях выгодно использовать один вид записи, а в некоторых — другой.

Мы будем опускать все доказательства сходимости, поскольку они следуют общему образцу, изложенному в гл. II.

2. Приложение квазилинеаризации к системам

Рассмотрим нелинейную систему

$$\frac{dx_i}{dt} = g_i(x_1, x_2, \ldots, x_N), \quad i = 1, 2, \ldots, N, \tag{2.1}$$

где x_i удовлетворяют ряду условий при $t=0$ и $t=b$.

Исключительно в целях простоты записи мы опускаем случай зависимости правых частей от t. Мы можем либо написать

$$\frac{dx_i}{dt} = g_i(x_1, x_2, \ldots, x_N, t), \quad i = 1, 2, \ldots, N, \tag{2.2}$$

либо ввести новую зависимую переменную x_{N+1} вместо t и написать

$$\begin{aligned} &\frac{dx_i}{dt} = g_i(x_1, x_2, \ldots, x_N, x_{N+1}), \quad i = 1, 2, \ldots, N, \\ &\frac{dx_{N+1}}{dt} = 1. \end{aligned} \tag{2.3}$$

Так как эта система того же вида, что и система (2.1), но размерности на единицу выше, можно ограничиться изучением системы (2.1).

Используя векторные обозначения, напишем

$$\frac{dx}{dt} = g(x), \tag{2.4}$$

где x есть N-мерный вектор с компонентами $x_1, x_2, \ldots, x_N$, а $g(x)$ есть N-мерный вектор с компонентами $g_i(x_1, x_2, \ldots, x_N)$, $i = 1, 2, \ldots, N$. Пусть краевые условия имеют вид

$$\begin{aligned} &(x(0), a_i) = b_i, \quad i = 1, 2, \ldots, k, \\ &(x(b), a_i) = b_i, \quad i = k+1, \ldots, N, \end{aligned} \tag{2.5}$$

где a_i — заданные векторы, b_i — заданные скаляры, а (x, y), как обычно, обозначает скалярное произведение

$$(x, y) = \sum_{i=1}^{N} x_i y_i. \tag{2.6}$$

Излагаемый ниже общий метод нетрудно использовать и в случае, когда заданы нелинейные краевые условия.

Как и раньше, без труда применяем квазилинеаризацию. Решая систему линейных уравнений

$$\frac{dx^{n+1}}{dt} = g(x^n) + J(x^n)(x^{n+1} - x^n), \quad n = 0, 1, 2, \ldots, \tag{2.7}$$

с граничными условиями

$$\begin{aligned} &(x^{n+1}(0), a_i) = b_i, \quad i = 1, 2, \ldots, k, \\ &(x^{n+1}(b), a_i) = b_i, \quad i = k+1, \ldots, N, \end{aligned} \tag{2.8}$$

и начальным приближением x^0, получаем последовательность векторов $\{x^n\}$. Здесь $J(x^n)$ — матрица Якоби, определяемая формулой

$$J(x^n) = \left(\frac{\partial g_i}{\partial x_j}(x_1^n, x_2^n, \ldots, x_N^n)\right) = \left(\frac{\partial g_i}{\partial x_j}\right)_{x=x^n}, \quad i, j = 1, 2, \ldots, N. \tag{2.9}$$

Прежде чем привести числовой пример, посмотрим, к чему сводится решение линейной системы.

3. Решение линейной системы

Рассмотрим уравнение

$$\frac{dx}{dt} = A(t)x \tag{3.1}$$

с краевыми условиями

$$\begin{aligned} &(x(0), a_i) = b_i, \quad i = 1, 2, \ldots, k, \\ &(x(b), a_i) = b_i, \quad i = k+1, \quad k+2, \ldots, N. \end{aligned} \tag{3.2}$$

Если $X(t)$ — решение матричного уравнения

$$\frac{dX}{dt} = A(t)X, \quad X(0) = I, \tag{3.3}$$

где I — единичная матрица, то общее решение уравнения (3.1) имеет вид

$$x = Xc, \tag{3.4}$$

где c — вектор. Подставляя (3.4) в (3.2), получаем систему линейных алгебраических уравнений относительно компонент вектора c:

$$\begin{aligned} &(X(0)c, a_i) = b_i, \quad i = 1, 2, \ldots, k, \\ &(X(b)c, a_i) = b_i, \quad i = k+1, \ldots, N, \end{aligned} \tag{3.5}$$

или

$$\begin{aligned}(c,\ a_i) &= b_i, \quad i = 1,\ 2,\ \ldots,\ k,\\ (c,\ X(b)'\,a_i) &= b_i, \quad i = k+1,\ \ldots,\ N,\end{aligned} \tag{3.6}$$

где X' — матрица, транспонированная к X.

Условие единственности решения этой системы, как и следовало ожидать, совпадает с условием на собственные значения. Один из способов обеспечения этого состоит в том, чтобы отрезок $[0,\ b]$ был достаточно мал.

Особенно подчеркнем, что редукция численного решения первоначальной задачи к численному решению системы линейных алгебраических уравнений никоим образом не исчерпывает всех проблем. Сразу же возникают важные вопросы, связанные с точностью и устойчивостью решения. Некоторые из них мы обсудим ниже. Кроме того, мы кратко опишем некоторые другие подходы, позволяющие избежать решения линейных алгебраических систем. Одним из них является динамическое программирование.

4. Числовой пример

В качестве примера использования метода квазилинеаризации для исследования уравнений высшего порядка рассмотрим численное решение уравнения

$$u^{(4)} + ke^{u} = 0 \tag{4.1}$$

с краевыми условиями

$$u(0) = u'(0) = 0, \quad u''(1) = u'''(1) = 0 \tag{4.2}$$

Используя рекуррентное соотношение

$$\begin{aligned}&u_{n+1}^{(4)} + ke^{u_n} + k(u_{n+1} - u_n)\,e^{u_n} = 0,\\ &u_{n+1}(0) = u'_{n+1}(0) = 0, \quad u''_{n+1}(1) = u'''_{n+1}(1) = 0,\end{aligned} \tag{4.3}$$

Таблица 4.1

x	$u_1(x)$	$u_2(x)$	$u_3(x)$	$u_4(x)$
0.0	0,0	0,0	0,0	0,0
0.20312	$-0{,}375572 \cdot 10^{-1}$	$-0{,}394922 \cdot 10^{-1}$	$-0{,}394976 \cdot 10^{-1}$	$-0{,}394976 \cdot 10^{-1}$
0.40325	$-0{,}128544$	$-0{,}135598$	$-0{,}135617$	$-0{,}135617$
0.60156	$-0{,}242775$	$-0{,}256792$	$-0{,}256831$	$-0{,}256831$
0.8469	$-0{,}373308$	$-0{,}395724$	$-0{,}395786$	$-0{,}395786$
1.0	$-0{,}501782$	$-0{,}532645$	$-0{,}532730$	$-0{,}532730$

с начальным приближением $u_0(x)=0$ и действуя в соответствии с вышеизложенным, мы получаем табл. 4.1 ($k=6$). Здесь применялся метод интегрирования Адамса — Мультона. Полный расчет на IBM-7090 потребовал около восьми секунд.

5. Многоточечные краевые задачи

Методика, использованная нами в предыдущем разделе, в равной степени применима в случае численного решения уравнения вида

$$\frac{dx}{dt}=g(x), \tag{5.1}$$

где условия заданы в нескольких точках, например

$$(x(t_i),\ a_i)=b_i, \qquad i=1,\ 2,\ \ldots,\ N, \tag{5.2}$$

причем $t_1\leqslant t_2\leqslant\ldots\leqslant t_N$. Задачи такого типа самым естественным образом возникают при идентификации систем.

В более общем случае можно рассматривать дополнительные условия вида

$$\int_a^b (x(t),\ dk_i(t))=b_i, \qquad i=1,\ 2,\ \ldots,\ N, \tag{5.3}$$

где интегралы понимаются в смысле Римана — Стильтьеса, а также и нелинейные варианты этих соотношений. В гл. VI будет приведено несколько примеров численного решения многоточечных краевых задач. Хотя эти задачи имеют важное научное значение, для их аналитического исследования сделано сравнительно мало.

6. Последовательные приближения и память

На предыдущих страницах мы рассматривали решение краевой задачи

$$\begin{aligned}&\frac{dx}{dt}=g(x),\\ &(x(0),\ a_i)=b_i, \qquad i=1,\ 2,\ \ldots,\ k,\\ &(x(b),\ a_i)=b_i, \qquad i=k+1,\ \ldots,\ N,\end{aligned} \tag{6.1}$$

с помощью последовательных приближений, порождаемых соотношением

$$\begin{aligned}&\frac{dx_{n+1}}{dt}=g(x_n)+J(x_n)(x_{n+1}-x_n),\\ &(x_{n+1}(0),\ a_i)=b_i, \qquad i=1,\ 2,\ \ldots,\ k,\\ &(x_{n+1}(b),\ a_i)=b_i, \qquad i=k+1,\ \ldots,\ N.\end{aligned} \tag{6.2}$$

Кроме того, мы указали, каким образом можно непосредственно использовать квазилинеаризацию для исследования вариационных задач.

Приступим теперь к фактическим вычислительным аспектам изложенного ранее алгоритма. Для того чтобы с помощью уравнения (6.2) получить значения $x_{n+1}(t)$ при заданных значениях $x_n(t)$ в узловых точках $t=0, \Delta, 2\Delta, \ldots, M\Delta$, необходимо хранить в памяти набор значений $\{x_n(k\Delta)\}$, $k=0, 1, 2, \ldots, M$. Это обычная ситуация, когда используется метод последовательных приближений, и мы уже встречались с ней в предыдущих расчетах.

Этот вопрос не был бы столь серьезен, если бы речь шла о решении одного уравнения типа, рассмотренного в разд. 4, или даже о решении системы невысокой размерности. Однако, если мы собираемся исследовать систему, размерность которой 50 или 100, и при этом желаем получить решение с определенной точностью, так что для каждой функции необходимо хранить около 1000 значений, то нехватка оперативной памяти может сделать невозможным предлагаемое нами решение.

В следующем разделе мы изложим метод, который до некоторой степени позволит нам преодолеть эту трудность. В связи с этим мы столкнемся с очень интересными и трудными задачами хранения и пересылки информации, которые абсолютно нигде не рассматривались.

7. Совместный расчет приближений

Рассмотрим общую задачу нахождения численного решения векторного уравнения

$$\frac{dx}{dt} = g(x) \tag{7.1}$$

с линейными краевыми условиями

$$\begin{aligned} (x(0),\ a_i) &= b_i, \qquad i = 1,\ 2,\ \ldots,\ k, \\ (x(b),\ a_i) &= b_i, \qquad i = k+1,\ \ldots,\ N. \end{aligned} \tag{7.2}$$

Как обычно, заменим уравнение (7.1) последовательностью уравнений

$$\frac{dx_{n+1}}{dt} = g(x_n) + J(x_n)(x_{n+1} - x_n), \tag{7.3}$$

где каждое x_n удовлетворяет условиям (7.2).

Пусть $x_0(t)$ — начальное приближение, которое не требует значительного объема памяти, скажем вектор, все компоненты которого являются либо константами, либо полиномами от t, либо экспонентами. Пусть $x_1(t)$ определяется из (7.2) и (7.3) при $n=0$.

Напомним, как находится x_1. Если матрица X_1 — решение уравнения

$$\frac{dX_1}{dt} = J(x_0)\,X_1, \quad X_1(0) = I, \tag{7.4}$$

и вектор p_1 — решение уравнения

$$\frac{dp_1}{dt} = J(x_0)\,p_1 + g(x_0) - J(x_0)\,x_0, \quad p_1(0) = 0, \tag{7.5}$$

то решение уравнения (7.3) имеет вид

$$x_1 = X_1 c_1 + p_1, \tag{7.6}$$

где вектор c_1 определяется из системы N линейных уравнений

$$\begin{aligned} (c_1,\ a_i) &= b_i, \qquad i = 1,\ 2,\ \ldots,\ k, \\ (X_1(b)\,c_1 + p_1(b),\ a_i) &= b_i, \qquad i = k+1,\ \ldots,\ N. \end{aligned} \tag{7.7}$$

Предполагая, что решение этой системы единственно (это имеет место, когда b достаточно мало), находим c_1. Отсюда получаем полный набор начальных условий

$$x_1(0) = X_1(0)\,c_1 + p_1(0) = c_1. \tag{7.8}$$

Для выполнения указанных операций необходимо совместно решить систему N^2+N дифференциальных уравнений первого порядка и систему N алгебраических уравнений.

Для нахождения x_2 необходимо решить систему

$$\begin{aligned} \frac{dX_2}{dt} &= J(x_1)\,X_2, \quad X_2(0) = I, \\ \frac{dp_2}{dt} &= J(x_1)\,p_2 + g(x_1) - J(x_1)\,x_1, \quad p_2(0) = 0, \end{aligned} \tag{7.9}$$

где вектор x_1 определен ранее. Вместо того чтобы хранить в памяти машины значения $x_1(t)$, дополним (7.9) уравнениями

$$\begin{aligned} \frac{dx_1}{dt} &= J(x_0)\,x_1 + g(x_0) - J(x_0)\,x_0, \\ x_1(0) &= c_1, \end{aligned} \tag{7.10}$$

где начальные значения определяются условиями (7.8).

Теперь необходимо решить совместно $(N^2+N)+N$ дифференциальных уравнений плюс N алгебраических уравнений для

определения постоянного вектора c_2, фигурирующего в представлении $x_2 = X_2 c_2 + p_2$. Вектор c_2 представляет собой полный набор начальных условий для определения x_2.

Продолжая действовать таким образом, видим, что на R-м шаге необходимо решить систему дифференциальных уравнений

$$\begin{aligned}
&\frac{dX_R}{dt} = J(x_{R-1})\,X_R, \quad X_R(0) = I,\\
&\frac{dp_R}{dt} = J(x_{R-1})\,p_R + g(x_{R-1}) - J(x_{R-1})\,x_{R-1}, \quad p_R(0) = 0, \qquad (7.11)\\
&\frac{dx_i}{dt} = J(x_{i-1})\,x_i + g(x_{i-1}) - J(x_{i-1})\,x_{i-1}, \quad x_i(0) = c_i,\\
&\qquad\qquad i = 1, 2, \ldots, R-1.
\end{aligned}$$

Это означает, что необходимо решить систему $N^2 + RN$ дифференциальных уравнений с заданными начальными условиями; при этом нет нужды запоминать предыдущие приближения. На каждом шаге мы должны решать систему N линейных алгебраических уравнений.

8. Обсуждение

Так как сходимость квадратичная, то можно ожидать, что R редко будет превышать 5 или самое большее 10. Отсюда, если $N=2$, то нам придется иметь дело самое большее с 24 совместными уравнениями; если $N=4$, то самое большее с 56; если $N=10$, то самое большее с 200 уравнениями. Эти требования весьма умеренны с точки зрения возможностей современных вычислительных машин.

С другой стороны, если N достаточно велико, скажем равно 100, то мы видоизменяем наш метод следующим образом. Вместо того чтобы решать матричное уравнение

$$\frac{dX_{n+1}}{dt} = J(x_n)\,X_{n+1}, \quad X_{n+1}(0) = I, \qquad (8.1)$$

в один прием (операция, которая требует интегрирования системы N^2 уравнений вместе с N уравнениями для определения x_n), мы будем последовательно определять столбцы матрицы X_{n+1} при помощи уравнений

$$\frac{dy_j}{dt} = J(x_n)\,y_j, \quad y_j(0) = e_j, \qquad (8.2)$$

где e_j — вектор, j-я компонента которого равна единице, а остальные нулю:

$$e_j = \begin{pmatrix} 0 \\ \cdot \\ \cdot \\ \cdot \\ 1 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{pmatrix}. \tag{8.3}$$

Это требует решения $N+RN$ совместных уравнений N раз. Как и обычно, за счет выигрыша в объеме памяти, которой у нас не всегда хватает, мы проигрываем во времени счета. Далее в связи с дифференциальной аппроксимацией мы рассмотрим другой метод, позволяющий уменьшить размерность системы дифференциальных уравнений.

9. Прямое и обратное интегрирование

Теперь мы кратко упомянем еще один подход к решению двухточечных краевых задач, позволяющий избежать как затруднений с памятью, так и решения систем линейных алгебраических уравнений.

Чтобы проиллюстрировать эту идею в простейшем виде, рассмотрим скалярное уравнение

$$u'' = g(u, u', t), \quad u(0) = u(b) = 0, \tag{9.1}$$

и поступим следующим образом. Зададимся некоторым значением $u'(0)$ и проинтегрируем численно (9.1) с начальными условиями $u(0)=0$, $u'(0)=c_1$. При этом мы получим некоторые значения $u(b)$, $u'(b)$. Если $u(b)=0$ с заданной степенью точности, то решение закончено. В противном случае мы интегрируем обратно от $t=b$ с начальными условиями $u(b)=0$, $u'(b)=d_1$, где d_1 — значение $u'(b)$, полученное при прямом интегрировании. При этом мы получаем некоторые значения $u(0)$, $u'(0)$. Если $u(0)=0$ с заданной степенью точности, то решение закончено. В противном случае интегрируем слева направо с начальными условиями $u(0)=0$, $u'(0)=c_2$, где c_2 — значение $u'(0)$, полученное при обратном интегрировании.

Можно показать, что в ряде случаев этот метод приводит к последовательности $\{c_n\}$, сходящейся к точному значению

$u'(0)$[1]). Чтобы расширить область применимости, потребуется более тонкий подход, основанный, возможно, на методе суммирования.

10. Дифференциально-разностные уравнения

Физические процессы с запаздыванием по времени описываются не традиционными дифференциальными уравнениями, а дифференциально-разностными уравнениями. Типичное уравнение имеет вид

$$u'(t) = g(u(t),\ u(t-1)),\quad t \geqslant 1, \tag{10.1}$$

где функция $u(t)$ теперь задана на начальном отрезке:

$$u(t) = u_0(t),\quad 0 \leqslant t \leqslant 1. \tag{10.2}$$

Использование любого из числа стандартных вычислительных методов, применимых к обыкновенным дифференциальным уравнениям, для получения численного решения уравнений такой природы не вызывает затруднений. Однако при этом мы вынуждены хранить в памяти значения функции на отрезке единичной длины. Можно ли избежать этого? Опять-таки это нетрудно, когда мы имеем дело с системами малой размерности, но в случае систем высокой размерности возникают затруднения в связи с нехваткой памяти.

11. Редукция к дифференциальным уравнениям

Начнем с того, что введем последовательность функций

$$u_n(t) = u(t+n),\quad 0 \leqslant t \leqslant 1,\quad n = 1,\ 2,\ \ldots. \tag{11.1}$$

Тогда одно дифференциально-разностное уравнение (10.1) можно переписать в виде системы уравнений

$$u_n'(t) = g(u_n(t),\ u_{n-1}(t)),\quad n = 1,\ 2,\ \ldots, \tag{11.2}$$

и заданной функции $u_0(t)$, а начальные условия таковы:

$$u_n(0) = u_{n-1}(1). \tag{11.3}$$

Существуют две трудности, которые нужно преодолеть, чтобы получить численное решение системы (11.2): во-первых, неизвестны начальные значения и, во-вторых, необходимо помнить функцию $u_0(t)$.

[1]) Можно привести целый ряд примеров, в которых изложенный метод не сходится (например, для уравнения $u''=0$). — *Прим. ред.*

Забудем на некоторое время, что нужно помнить функцию $u_0(t)$ (мы ниже вернемся к этому вопросу), и покажем, как можно определить неизвестные значения $u_n(0)$. Рассмотрим уравнение

$$u_1'(t) = g(u_1(t),\ u_0(t)),\quad u_1(0) = u_0(1). \tag{11.4}$$

Интегрируя его численно, получаем значение $u_1(1) = u_2(0)$. Теперь повторим все снова и проинтегрируем систему

$$\begin{aligned} u_1'(t) &= g(u_1(t),\ u_0(t)),\quad u_1(0) = u_0(1),\\ u_2'(t) &= g(u_2(t),\ u_1(t)),\quad u_2(0) = u_1(1), \end{aligned} \tag{11.5}$$

откуда получаем $u_2(1) = u_3(0)$.

Продолжая далее действовать так же, мы на k-м шаге разрешаем систему, состоящую из k уравнений

$$u_i'(t) = g(u_i(t),\ u_{i-1}(t)),\quad u_i(0) = b_i,\qquad i = 1,\ 2,\ \ldots,\ k, \tag{11.6}$$

где b_i — известные величины.

Таким образом, для определения решения на всем отрезке $[0,\ N]$ необходимо проинтегрировать в общей сложности $1+2+\ldots+N = N(N+1)/2$ уравнений. Если N не слишком велико, то это довольно стандартная операция, требующая минимального объема памяти и занимающая не так много времени. Как мы увидим в разделе, посвященном дифференциальной аппроксимации, в нашем распоряжении имеются способы уменьшения затрачиваемого труда.

12. Функционально-дифференциальные уравнения

Применим тот же общий подход к функционально-дифференциальному уравнению

$$u'(t) = g(t,\ u(t),\ u(h(t))). \tag{12.1}$$

Уравнения такого типа возникают при математическом моделировании в целом ряде областей, начиная от электромагнитной теории и теории управления и кончая физиологией дыхания и нейрофизиологией.

Предположим, что $h(t) \leqslant t$ при $t \geqslant 0$, так что будущее определяется прошлым. Кроме того, предположим, что $h'(t) > 0$ при $t \geqslant 0$, и обозначим $h^{-1}(t)$ через $H(t)$. Допустим, что функция $u(t)$ известна на некотором начальном отрезке $[0,\ t_1]$, где $t_1 = H(0)$.

Определим последовательность $\{t_n\}$ рекуррентными соотношениями

$$t_n = H(t_{n-1}),\qquad n = 2,\ 3,\ \ldots. \tag{12.2}$$

Пусть $H^{(n)}(t)$ обозначает n-ю итерацию $H(t)$, $H^{(n)}=H(H^{(n-1)})$, $n=2, 3, \ldots$. В силу наших предположений функция $H(t)$ взаимно однозначно отображает отрезок $[t_{n-1}, t_n]$ на $[t_n, t_{n+1}]$ и $H^{(k)}(t)$ отображает $[t_{n-1}, t_n]$ на $[t_{n-1+k}, t_{n+k}]$.

Рассмотрим функцию

$$u_n(s)=u(H^{(n)}(s)), \qquad n=0, 1, 2, \ldots, \tag{12.3}$$

где s лежит на отрезке $[0, t_1]$ и $H^{(0)}(s)=s$. Тогда значения $u_n(s)$ при $0\leqslant s\leqslant t_1$ совпадают со значениями $u(t)$ при $t_n\leqslant t\leqslant t_{n+1}$.

Мы имеем

$$u_n'(s)=\left[\frac{d}{ds}H^{(n)}(s)\right]u'(H^{(n)}(s)), \tag{12.4}$$

а производную $H^{(n)}(s)$ можно легко вычислить при помощи рекуррентного соотношения

$$\frac{d}{ds}H^{(n)}(s)=H'(H^{(n-1)}(s))\frac{d}{ds}H^{(n-1)}(s). \tag{12.5}$$

Положим теперь $t=H^{(n)}(s)$, где $0\leqslant s\leqslant t_1$. Тогда из уравнения (12.1) получаем [1])

$$u'(H^{(n)}(s))=g\{u(H^{(n)}(s)),\ u(H^{(n-1)}(s))\}=$$
$$=g(u_n(s),\ u_{n-1}(s)), \qquad n=1, 2, \ldots. \tag{12.6}$$

Учитывая (12.4), равенство (12.6) можно переписать в виде

$$u_n'(s)=\left[\frac{d}{ds}H^{(n)}(s)\right]g(u_n(s),\ u_{n-1}(s)), \qquad n=1, 2, \ldots. \tag{12.7}$$

Итак, уравнение (12.1) заменено системой обыкновенных дифференциальных уравнений, где s изменяется в фиксированном отрезке $[0, t_1]$. Теперь можно приступить к нахождению численного решения с помощью итерационной схемы разд. 11.

13. Дифференциальная аппроксимация

Если мы хотим получить решение на большом интервале времени, то предыдущий метод может оказаться громоздким из-за большого числа дифференциальных уравнений, которые необходимо решить совместно. Нам хотелось бы развить метод, при использовании которого в каждый момент времени $t=k$ в качестве новой начальной функции выбиралась бы функция $u_k(t)$. Затруднение состоит в том, что мы теперь должны запоминать значения функции, чего мы хотели избежать вначале.

[1]) Функция g считается здесь не зависящей явно от t. — *Прим. ред.*

Один из возможных путей преодоления этого затруднения состоит в аппроксимации функции $u_k(t)$ функцией простой аналитической структуры, такой, как полином, экспоненциальный полином или кусочно-линейная функция. Так, например, если мы аппроксимируем $u_k(t)$ на отрезке $[0, T]$ при помощи полинома

$$u_k(t) \cong \sum_{n=0}^{M} a_n t^n, \tag{13.1}$$

то нам необходимо помнить только $M+1$ коэффициентов $[a_0, a_1, \ldots, a_M]$ и правило построения полинома по заданным коэффициентам. То же замечание справедливо, если мы используем аппроксимацию вида

$$u_k(t) \cong \sum_{n=1}^{M} a_n e^{\lambda_n t}. \tag{13.2}$$

Обе эти аппроксимации являются частными случаями приближения функции $u(t)$ решением v линейного дифференциального уравнения

$$v^{(N)} + a_1 v^{(N-1)} + \ldots + a_N v = 0, \quad v^{(i)}(0) = c_i, \quad i = 0,1, \ldots, N-1. \tag{13.3}$$

Один из путей построения такой аппроксимации состоит в нахождении величин a_i и c_i, доставляющих минимум квадратичному функционалу

$$\int_0^T (v-u)^2 dt. \tag{13.4}$$

Эту нелинейную задачу оптимизации (частный случай общей проблемы управления и проектирования) можно исследовать при помощи квазилинеаризации. Этот вопрос будет обсужден в гл. VI. Здесь мы изложим более простой метод аппроксимации, который часто приводит к удовлетворительным результатам.

14. Простой вариант дифференциальной аппроксимации

Предположим, что $u(t)$ удовлетворяет обыкновенному дифференциальному уравнению

$$u^{(m)} = h(u, u', \ldots, u^{(m-1)}, t), \quad u^{(i)}(0) = c_i, \quad i = 0, 1, \ldots, m-1, \tag{14.1}$$

и мы хотим найти аппроксимирующее линейное дифференциальное уравнение вида

$$v^{(N)} + a_1 v^{(N-1)} + \ldots + a_N v = 0. \tag{14.2}$$

Определим коэффициенты a_i из условия, что они минимизируют квадратичное выражение

$$\int_0^T (u^{(N)} + a_1 u^{(N-1)} + \ldots + a_N u)^2 \, dt, \tag{14.3}$$

где функция u удовлетворяет уравнению (14.1).

Минимизация выражения (14.3) приводит к системе N линейных уравнений

$$\int_0^T u^{(j)} (u^{(N)} + a_1 u^{(N-1)} + \ldots + a_N u) \, dt = 0, \quad j = 0,1,2, \ldots, N-1. \tag{14.4}$$

Для умеренных значений N, т. е. для $N \leqslant 20$, численное решение не вызывает затруднений, как только мы вычислим интегралы, входящие в качестве коэффициентов. При желании мы могли бы произвести интегрирование по частям и свести вычисление этих интегралов к вычислению интегралов $\int_0^T (u^{(i)})^2 \, dt$. Однако для умеренных значений N более удобно поступить так, как это изложено далее.

Введем новые переменные u_{ij}, $i, j = 0, 1, \ldots, N-1$, определяемые уравнениями

$$\frac{du_{ij}}{dt} = u^{(N-i)} u^{(j)}, \quad u_{ij}(0) = 0, \tag{14.5}$$

и решим эти дифференциальные уравнения вместе с первоначальным уравнением для u:

$$u^{(N)} = g(u, u', \ldots, u^{(N-1)}, t). \tag{14.6}$$

Величины $u_{ij}(T)$ являются искомыми коэффициентами в (14.4).

Определив коэффициенты a_i посредством предыдущих процедур, мы хотим теперь найти аппроксимирующую функцию $v(t)$ как решение уравнения (14.2). Первая мысль, которая приходит на ум, — использовать начальные значения

$$v^i(0) = u^{(i)}(0), \qquad i = 0, 1, 2, \ldots, N-1, \tag{14.7}$$

и, действительно, ниже мы довольно успешно проделаем это.

Однако мы обычно поступаем следующим образом. Пусть $v_1, v_2, \ldots, v_N$ — фундаментальная система решений уравнения (14.2), определенная при условии, что матрица, столбцами которой являются $(v_1(0), v_1'(0), \ldots, v_1^{(N-1)}(0))$ и т. д., есть единичная матрица.

Тогда любое решение можно записать в виде

$$u(t)=\sum_{i=1}^{N} c_i v_i, \tag{14.8}$$

где c_i — скаляры. Выберем c_i так, чтобы достигался минимум выражения

$$\int_0^T \left(u-\sum_{i=1}^{N} c_i v_i\right)^2 dt. \tag{14.9}$$

Уравнения для c_i имеют вид

$$\int_0^T u v_i\,dt-\sum_{j=1}^{N} c_j \int_0^T v_i v_j\,dt=0, \qquad i=1,\ 2,\ \ldots,\ N. \tag{14.10}$$

Для вычисления интегралов введем, как и раньше, переменные w_i и w_{ij} при помощи соотношений

$$\begin{aligned} \frac{dw_i}{dt}&=uv_i, \qquad v_i(0)=0,\\ \frac{dw_{ij}}{dt}&=v_iv_j, \quad w_{ij}(0)=0, \end{aligned} \tag{14.11}$$

добавим к ним уравнения для v_i (уравнения (14.2) с соответствующими граничными условиями) и уравнение для u и проинтегрируем полученную систему.

15. Уравнение восстановления

Рассмотрим теперь интегральное уравнение типа

$$u(t)=f(t)+\int_0^t e^{-(t-s)^2}u(s)\,ds. \tag{15.1}$$

При $0\leqslant t\leqslant 1$ мы получаем в качестве дифференциальной аппроксимации третьего порядка для функции e^{-t^2} решение уравнения

$$u^{(3)}+2{,}740299u^{(2)}+7{,}9511452u^{(1)}+5{,}7636455u=0. \tag{15.2}$$

Используя начальные условия, полученные при помощи функции e^{-t^2}, а именно

$$u(0)=1, \quad u'(0)=0, \quad u''(0)=-2, \tag{15.3}$$

мы получаем на всем отрезке $0\leqslant t\leqslant 1$ настолько хорошее совпадение между e^{-t^2} и решением $u(t)$ уравнения (15.2), что нет нужды использовать методику разд. 14.

Рассмотрим выражение

$$w(t)=\int_0^t k(t-s)\,u(s)\,ds. \tag{15.4}$$

Дифференцируя его несколько раз и используя коэффициенты, полученные выше, имеем

$$\begin{aligned} w^{(3)}+2{,}740299w^{(2)}+7{,}9511452w^{(1)}+5{,}7636455w=\\ =k(0)\,u''(t)+k'(0)\,u'(t)+k''(0)\,u(t)+\\ +2{,}740299\,[k(0)\,u'(t)+k'(0)\,u(t)]+7{,}9511452k(0)\,u(t)+\\ +\int_0^t u(s)\,[k'''(t-s)+2{,}740299k''(t-s)+\\ +7{,}9511452k'(t-s)+5{,}7636455k(t-s)]\,ds. \end{aligned} \tag{15.5}$$

Выбирая $k(t)=e^{-t^2}$ и предполагая, что членом, стоящим под знаком интеграла, можно пренебречь, получаем линейное дифференциальное уравнение третьего порядка для $w=u-f$.

Возьмем $f=1-\int_0^t e^{-s^2}\,ds$, так что уравнение

$$u(t)=1-\int_0^t e^{-s^2}\,ds+\int_0^t e^{-(t-s)^2}u(s)\,ds \tag{15.6}$$

имеет решение $u(t)=1$.

Выбранная функция $f(t)$ удовлетворяет линейному дифференциальному уравнению

$$f^{(3)}+2tf^{(2)}+2f^{(1)}=0 \tag{15.7}$$

Таблица 4.2

t	$u(t)$	$u'(t)$	$u''(t)$
0,1	0,999999	—	—
0,2	0,999999	$-0{,}146\cdot10^{-3}$	$-0{,}148\cdot10^{-2}$
0,3	0,999969	—	—
0,4	0,999937	—	—
0,5	0,999909	—	—
0,6	0,999898	—	—
0,7	0,999909	$0{,}299\cdot10^{-3}$	$0{,}174\cdot10^{-3}$
0,8	0,999938	$0{,}330\cdot10^{-3}$	$0{,}167\cdot17^{-3}$
0,9	0,999970	$0{,}272\cdot10^{-3}$	$0{,}135\cdot10^{-2}$
1,0	0,999989	$0{,}919\cdot10^{-4}$	$0{,}189\cdot10^{-2}$

при начальных условиях

$$f(0)=1, \quad f'(0)=-1, \quad f''(0)=0.$$

Решая (15.7) вместе с приближенным линейным уравнением для u, полученным из (15.5), находим значения $u(t)$, приведенные в табл. 4.2.

Из таблицы видно превосходное совпадение приближенного решения с точным решением $u(t)=1$.

16. Обсуждение

Во многом ту же методику можно использовать, когда $u(t)$ удовлетворяет дифференциально-разностному уравнению. С другой стороны, можно задать набор значений t_i и поставить задачу о минимизации выражения

$$\sum_{i=1}^{M} (u(t_i)-v(t_i))^2 \tag{16.1}$$

путем соответствующего выбора коэффициентов и начальных условий. Мы подробно разберем этот вопрос в гл. VI в связи с некоторыми проблемами идентификации.

Отметим также, что комбинацию дифференциальной аппроксимации и квазилинеаризации можно использовать для того, чтобы упростить исследование интегро-дифференциального уравнения вида

$$\begin{gathered} u''+g(u)+\int_0^1 k(x-x_1)\,h(u)\,dx_1+r(x)=0, \\ u(0)=u(1)=0. \end{gathered} \tag{16.2}$$

Предположим, как это часто бывает, что $k(x)=k(-x)$, и примем для простоты, что решение линейных уравнений дифференциальной аппроксимации порождает аппроксимацию вида

$$k(x)\cong\sum_{k=1}^{M}\alpha_k e^{\lambda_k x}, \quad 0\leqslant x\leqslant 1 \tag{16.3}$$

(считаем, что корни простые). Тогда

$$\begin{aligned} \int_0^1 k(x-x_1)\,h(u)\,dx_1 &= \sum_{k=1}^{M}\int_0^1 \alpha_k e^{\lambda_k|x-x_1|}h(u)\,dx_1 = \\ &= \sum_{k=1}^{M}\alpha_k e^{\lambda_k x}\int_0^x e^{-\lambda_k x_1}h(u)\,dx_1 + \\ &\quad + \sum_{k=1}^{M}\alpha_k e^{-\lambda_k x}\int_x^1 e^{\lambda_k x_1}h(u)\,dx_1. \end{aligned} \tag{16.4}$$

Пусть

$$v_k = e^{\lambda_k x} \int_0^x e^{-\lambda_k x_1} h(u)\, dx_1,$$
$$w_k = e^{-\lambda_k x} \int_x^1 e^{\lambda_k x_1} h(u)\, dx_1. \tag{16.5}$$

Тогда

$$v_k' = \lambda_k v_k + h(u), \qquad v_k(0) = 0,$$
$$w_k' = -\lambda_k w_k - h(u), \qquad w_k(1) = 0, \tag{16.6}$$

и

$$u'' + g(u) + \sum_{k=1}^{M} \alpha_k (v_k + w_k) + r(x) = 0, \quad u(0) = u(1) = 0. \tag{16.7}$$

Теперь для получения решения этой двухточечной краевой задачи можно использовать стандартную методику квазилинеаризации.

Здесь читатель вправе задать вопрос, почему мы начали не с аппроксимации типа (16.3), а с более сложного приближения (13.3). В качестве ответа сошлемся на общеизвестные трудности, связанные с определением коэффициентов и показателей степеней в (16.3). Более точно уже простые примеры показывают, что λ_k являются неустойчивыми функционалами от $k(x)$. По существу это обусловлено влиянием на аппроксимацию кратных корней характеристического уравнения для (13.3)

$$\lambda^N + a_1 \lambda^{N-1} + \ldots + a_N = 0. \tag{16.8}$$

С другой стороны, a_i, будучи элементарными симметрическими функциями, являются устойчивыми функционалами. Следовательно, лучше попытаться сначала найти a_i, а затем переходить к решению дифференциального уравнения.

17. Запоминание и память

Предыдущее обсуждение совершенно ясно показывает, что мы должны быть очень точны, когда говорим о требованиях к запоминанию при численном решении разного типа проблем.

Всегда подразумевается, но, к сожалению, не всегда достаточно явно подчеркивается тот факт, что для каждого частного метода требуется свой определенный объем памяти. Более искусный метод может существенно уменьшить эти потребности.

Этот вопрос имеет непосредственное отношение к раскрытию тайн человеческой памяти. Когда отмечают, что для совершения

разнообразных действий человеческому мозгу требуется большое количество энергии или огромное число ячеек определенных типов, обычно полагают, что эти требования являются очевидной необходимостью в свете тех немногих процессов, которые мы в настоящее время понимаем или думаем, что понимаем. Вполне возможно и правдоподобно, что в эволюционном процессе проб и ошибок живые организмы разработали множество более изощренных способов для запоминания и извлечения информации, для распознавания образов и принятия решения. Открытие этих способов было бы выдающимся достижением, и, конечно, весьма самонадеянно предполагать, что усилия небольшой группы исследователей разрешат за короткий срок хотя бы некоторые загадки человеческого ума.

18. Свойство монотонности для систем

Рассмотрим теперь возможность получения некоторых результатов, аналогичных результатам гл. III, для систем дифференциальных неравенств вида

$$\frac{dx}{dt} \geqslant A(t)x, \quad x(0)=c, \tag{18.1}$$

где x есть N-мерный вектор. Сокращенная запись $x \geqslant y$ (x, y — векторы) здесь означает, что $x_i \geqslant y_i$, $i=1, 2, \ldots, N$. Мы хотим наложить простые условия на $A(t)$, которые гарантировали бы, что $x \geqslant y$ при $t \geqslant 0$, где y определяется из уравнения

$$\frac{dy}{dt} = A(t)y, \quad y(0)=c. \tag{18.2}$$

Если $A(t)=A$ — постоянная матрица, то можно легко показать, что для этого необходимо и достаточно, чтобы

$$a_{ij} \geqslant 0, \quad i \neq j. \tag{18.3}$$

Необходимость следует из разложения

$$e^{At} = I + At + O(t^2), \tag{18.4}$$

в чем мы убеждаемся, исследуя правую часть в непосредственной окрестности точки $t=0$.

Докажем достаточность для более общего случая переменных коэффициентов при помощи следующего простого рассуждения. Сделаем преобразование переменных

$$y_i = e^{\int_0^t a_{ii}(s)\,ds} z_i, \qquad i=1, 2, \ldots, N, \tag{18.5}$$

где y_i — компоненты вектора y, а z_i — компоненты нового вектора z. Тогда (18.2) примет вид

$$\frac{dz}{dt} = B(t)z, \quad z(0) = c, \tag{18.6}$$

где

$$B(t) = (b_{ij}(t)), \tag{18.7}$$

причем

$$b_{ij}(t, = \begin{cases} 0, & i = j. \\ a_{ij}(t)\, e^{\int_0^t (a_{jj} - a_{ii})\, ds}, & i \neq j. \end{cases} \tag{18.8}$$

Если $a_{ij}(t) \geqslant 0$, то все компоненты $b_{ij}(t)$ неотрицательны, и тем самым $z(t) \geqslant z(0)$ при $t \geqslant 0$. Отсюда, если $c \geqslant 0$, то $y(t) \geqslant 0$ при $t \geqslant 0$.

Так как (18.1) можно переписать в виде

$$\frac{dx}{dt} = A(t)x + f(t), \quad x(0) = c, \tag{18.9}$$

где $f(t) \geqslant 0$, мы видим, что $z = x - y$ удовлетворяет уравнению

$$\frac{dz}{dt} = A(t)z + f(t), \quad z(0) = 0. \tag{18.10}$$

То же рассуждение, что и выше (формулы (18.5)—(18.8)), показывает, что $z \geqslant 0$ при $t \geqslant 0$.

Результаты такого типа чрезвычайно полезны при исследовании системы уравнений

$$\frac{dx}{dt} = \max_q [A(q, t)x + b(q)], \tag{18.11}$$

играющей основную роль при изучении марковских процессов решения — важного класса процессов динамического программирования в стохастическом случае.

19. Свойство монотонности для линейных дифференциальных уравнений N-го порядка

Рассмотрим линейное дифференциальное выражение

$$L(u) = u^{(N)} + p_1(t)\, u^{(N-1)} + p_2(t)\, u^{(N-2)} + \ldots + p_N(t)\, u. \tag{19.1}$$

Обозначим через W_n вронскиан n-го порядка

$$W(f_1, f_2, \ldots, f_n) = \begin{vmatrix} f_1(t) & f_1'(t) & \ldots & f_1^{(n-1)}(t) \\ f_2(t) & f_2'(t) & \ldots & f_2^{(n-1)}(t) \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ f_n(t) & f_n'(t) & \ldots & f_n^{(n-1)}(t) \end{vmatrix}. \tag{19.2}$$

Тогда оказывается справедливым разбиение, аналогичное тому, которое использовалось в гл. III, разд. 4, для линейного дифференциального выражения второго порядка, а именно

$$L(u)=\frac{W_N}{W_{N-1}}\frac{d}{dt}\left(\frac{W_{N-1}^2}{W_{N-2}W_N}\cdots\right.$$
$$\left.\cdots\frac{d}{dt}\left(\frac{W_2^2}{W_1W_3}\left(\frac{d}{dt}\left(\frac{W_1^2}{W_0W_2}\frac{d}{dt}\left(\frac{u}{W_1}\right)\cdots\right)\right)\right)\right), \quad (19.3)$$

где $W_0=1$, а $u_1, u_2, \ldots, u_N$ образуют набор линейно независимых решений уравнения $L(u)=0$.

Введем теперь понятие, которое Пойа называет «свойством W». Набор решений $u_1, u_2, \ldots, u_{N-1}$ обладает свойством W в открытом интервале $(0, b)$, если в этом интервале

$$u_1>0, \quad W(u_1, u_2)>0, \ldots, \quad W(u_1, u_2, \ldots, u_{N-1})>0. \quad (19.4)$$

Легко убедиться с помощью разбиения (19.3), что при наличии свойства W из условий

$$L(u)\geqslant 0, \quad L(v)=0, \quad (19.5)$$

где $u^{(i)}(0)=v^{(i)}(0)$, $i=1, 2, \ldots, N-1$, следует $u\geqslant v$ на $(0, b)$. Это дальнейшее обобщение результата С. А. Чаплыгина, установленного в разд. 5 гл. III.

20. Обсуждение

Мы получили несколько результатов, позволяющих установить, когда последовательность приближений сходится монотонно. Как и следовало ожидать, эти результаты не столь изящны или полезны, как соответствующие результаты для уравнения второго порядка.

В случае уравнения высшего порядка нас в первую очередь интересуют вопросы квадратичной сходимости и точности. В следующем разделе мы обсудим метод, который оказался весьма полезным при получении численного решения линейного дифференциального уравнения с достаточной степенью точности. Трудно преувеличить его значение для эффективных приложений квазилинеаризации.

21. Использование метода ортогонализации Грама — Шмидта для численного решения линейных краевых задач

На пути получения численного решения систем линейных алгебраических уравнений нас подстерегает много ловушек. Не существует метода решения таких систем, который можно было

бы считать полностью удовлетворительным. В одном из предыдущих разделов мы показали, как использовать метод прямого и обратного интегрирования, чтобы избежать численного решения линейных алгебраических систем в процессе решения линейной двухточечной краевой задачи. С другой стороны, можно выбрать частные решения однородного дифференциального уравнения таким образом, чтобы получить решения линейных уравнений для множителей с максимально возможной точностью. В этом разделе мы обсудим один из вариантов, в котором используется метод ортогонализации Грама — Шмидта.

Рассмотрим дифференциальное уравнение четвертого порядка

$$u^{(4)} - 24u^{(3)} - 169u^{(2)} - 324u^{(1)} - 180u = 0. \tag{21.1}$$

Его общее решение имеет вид

$$u(t) = s_1 e^{-t} + s_2 e^{-2t} + s_3 e^{-3t} + s_4 e^{+30t}. \tag{21.2}$$

Обратим внимание на частное решение

$$u(t) = e^{-t} - 2e^{-2t} + e^{-3t}, \tag{21.3}$$

которое удовлетворяет начальным условиям

$$\begin{aligned} u(0) &= 0, \\ \dot{u}(0) &= 0, \\ u^{(2)}(0) &= 2, \\ u^{(3)}(0) &= -12{,}0. \end{aligned} \tag{21.4}$$

При $t=1$ это решение удовлетворяет еще условиям

$$\begin{aligned} u(1) &= e^{-1} - 2e^{-2} + e^{-3} = 0{,}146996, \\ \dot{u}(1) &= -e^{-1} + 4e^{-2} - 3e^{-3} = 0{,}241005 \times 10^{-1}. \end{aligned} \tag{21.5}$$

Попытаемся определить функцию $u(t)$, задаваемую соотношением (21.3), как решение уравнения (21.1) при краевых условиях $u(0) = 0$, $\dot{u}(0) = 0$, $u(1) = 0{,}146996$, $\dot{u}(1) = 0{,}0241005$.

Если мы непосредственно представим решение краевой задачи в виде линейной комбинации частных решений и воспользуемся методом интегрирования Рунге — Кутта с шагом интегрирования 0,01, то придем к неутешительным результатам. Пусть $w(t)$ — решение уравнения (21.1) с начальными условиями

$$w(0) = 0, \quad \dot{w}(0) = 0, \quad w^{(2)}(0) = 1, \quad w^{(3)}(0) = 0, \tag{21.6}$$

и $z(t)$ — решение, для которого

$$z(0) = 0, \quad \dot{z}(0) = 0, \quad z^{(2)}(0) = 0, \quad z^{(3)}(0) = 1. \tag{21.7}$$

Численное интегрирование дает значения

$$\begin{aligned} w(1) &= 0{,}195694 \times 10^{10}, \\ \dot{w}(1) &= 0{,}587083 \times 10^{11}, \\ z(1) &= 0{,}326157 \times 10^{9}, \\ \dot{z}(1) &= 0{,}978472 \times 10^{10}. \end{aligned} \tag{21.8}$$

Используя эти значения при решении линейных алгебраических уравнений

$$\begin{aligned} w(1)\,c_1 + z(1)\,c_2 &= u(1), \\ \dot{w}(1)\,c_1 + \dot{z}(1)\,c_2 &= \dot{u}(1), \end{aligned} \tag{21.9}$$

находим

$$\begin{aligned} c_1 &= -0{,}609136 \times 10^{-2}, \\ c_2 &= 0{,}365481 \times 10^{-1}, \end{aligned} \tag{21.10}$$

причем $c_1 = u^{(2)}(0)$, $c_2 = u^{(3)}(0)$ [1]).

Так как точные значения, согласно (21.4), равны $u^{(2)}(0) = 2$, $u^{(3)}(0) = -12$, мы видим, что задача отыскания начальных значений функции $u(t)$ нами фактически не решена.

Трудность связана с тем, что собственные значения, соответствующие дифференциальному оператору (21.2), т. е. числа -1, -2, -3, 30, сильно отличаются друг от друга. Это означает, что одно из частных решений будет превалировать над другими, и, следовательно, возникает затруднение с определением линейной комбинации частных решений, удовлетворяющей заданным краевым условиям. Мы сталкиваемся с явлением «плохой обусловленности» [2]).

Чтобы исправить положение, мы уделим больше внимания избираемым частным решениям. Рассмотрим векторы

$$\alpha = \begin{pmatrix} w \\ \dot{w} \\ w^{(2)} \\ w^{(3)} \end{pmatrix}, \quad \beta = \begin{pmatrix} z \\ \dot{z} \\ z^{(2)} \\ z^{(3)} \end{pmatrix}, \tag{21.11}$$

---

[1]) Последние равенства следуют из того, что приближенное решение ищется в виде $u = c_1 w + c_2 z$, где w и z подчинены условиям (21.6) и (21.7). — *Прим. ред.*

[2]) Определитель линейной системы (21.9) оказывается малым по сравнению с характерной величиной его элементов, определяемых равенствами (21.8). — *Прим. ред.*

где функции w, z и их производные определены выше, т. е. w и z являются решениями однородного уравнения (21.1) с начальными условиями (21.6) и (21.7). Используя метод интегрирования Рунге — Кутта, сделаем пять шагов величины 0,01 и получим значения векторов $\alpha(0{,}05)$ и $\beta(0{,}05)$. Затем выполним ортогонализацию Грама — Шмидта над этими векторами, приводящую к новым ортонормированным векторам

$$\gamma(0{,}05) = \frac{\alpha(0{,}05)}{|\alpha(0{,}05)|} \tag{21.12}$$

и

$$\delta(0{,}05) = \frac{\beta(0{,}05) - (\beta(0{,}05),\ \gamma(0{,}05))\,\gamma(0{,}05)}{|\beta(0{,}05) - (\beta(0{,}05),\ \gamma(0{,}05))\,\gamma(0{,}05)|}, \tag{21.13}$$

где (α, β) и $|\alpha|$ обозначают, как обычно, скалярное произведение и длину вектора. Производя то же линейное преобразование над начальными векторами (21.6) и (21.7), т. е.

$$\begin{pmatrix}0\\0\\1\\0\end{pmatrix}, \quad \begin{pmatrix}0\\0\\0\\1\end{pmatrix},$$

находим новые начальные векторы, использование которых при интегрировании уравнения (21.1) приводит к ортогональным и нормированным векторам $\gamma(0{,}05)$ и $\delta(0{,}05)$. Эти начальные условия таковы:

$$\begin{pmatrix}0\\0\\1\\0\end{pmatrix} \frac{1}{|\alpha(0{,}05)|} \tag{21.14}$$

и

$$\frac{1}{|\beta(0{,}05) - (\beta(0{,}05),\ \gamma(0{,}05))\,\gamma(0{,}05)|} \left\{ \begin{pmatrix}0\\0\\0\\1\end{pmatrix} - (\beta(0{,}05),\ \gamma(0{,}05)) \begin{pmatrix}0\\0\\1\\0\end{pmatrix} \right\}. \tag{21.15}$$

Затем мы интегрируем уравнение (21.1) от $t=0{,}05$ до $t=0{,}10$, используя $\gamma(0{,}05)$ и $\delta(0{,}05)$ в качестве начальных условий при $t=0{,}05$. Полученные таким образом векторы ортогонализуются и соответствующее преобразование выполняется над начальными

условиями при $t=0$. Продолжая дальше этот процесс, мы находим в конце концов два вектора вида

$$\begin{pmatrix} 0 \\ 0 \\ a_3 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 0 \\ b_3 \\ b_4 \end{pmatrix},$$

которые, будучи выбранными в качестве начальных условий для уравнения (21.1), приводят к двум ортонормированным векторам при $t=1$. Таким образом, точно определено двумерное пространство при $t=1$, натянутое на решения уравнения (21.1), для которых $u(0)=\dot{u}(0)=0$. В некоторых случаях это помогает с высокой точностью определить недостающие условия $u^{(2)}(0)$ и $u^{(3)}(0)$.

Применяя этот метод к задаче (21.1) с прежними краевыми условиями $u(0)=0$, $u'(0)=0$, $u(1)=0{,}146996$, $u'(1)=0{,}0241005$, находим

$$\begin{aligned} u^{(2)}(0) &= 1{,}9999997, \\ u^{(3)}(0) &= -12{,}000003. \end{aligned} \tag{21.16}$$

Для уравнения

$$u^{(4)} - 9u^{(3)} - 79u^{(2)} - 159\dot{u} - 90u = 0, \tag{21.17}$$

общее решение которого имеет вид

$$u(t) = s_1 e^{-t} + s_2 e^{-2t} + s_3 e^{-3t} + s_4 e^{15t}, \tag{21.18}$$

использование упомянутого выше простого представления решения краевой задачи в виде линейной комбинации частных решений приводит к результатам $u^{(2)}(0)=2{,}0013081$, $u^{(3)}(0)=-12{,}007849$. Граничные условия здесь принимались такими же, как выше, и решение снова имеет вид (21.3). После двадцатикратного использования процесса Грама — Шмидта на отрезке $[0, 1]$ мы получаем $u^{(2)}(0)=1{,}9999995$, $u^{(3)}(0)=-12{,}000003$. В этом случае точные результаты по-прежнему равны $u^{(2)}(0)=2$, $u^{(3)}(0)=-12$.

22. Динамическое программирование

Укажем теперь, каким образом при исследовании задач минимизации квадратичных функционалов использовать динамическое программирование для того, чтобы избежать двухточечных краевых задач и решения сопутствующих им линейных алгебраических систем. Мы уже обсуждали одномерный

вариант такой задачи в предыдущей главе. Рассмотрим здесь векторный вариант.

Чтобы сделать выкладки по возможности проще, предположим, что мы хотим минимизировать функционал

$$J(x)=\int_0^T [(x',\ x')+(x,\ Ax)]\,dt \tag{22.1}$$

при условии $x(0)=c$. Пусть A — симметричная положительно определенная матрица, так что проблемы собственных значений не возникает.

Уравнение Эйлера и краевые условия порождают двухточечную краевую задачу [1])

$$x''-Ax=0,\quad x(0)=c,\quad x'(T)=0. \tag{22.2}$$

Поступим совсем иным образом, чем раньше. Запишем [2])

$$f(c,\ T)=\min_x J(x). \tag{22.3}$$

Тогда обычная методика динамического программирования приводит к нелинейному дифференциальному уравнению в частных производных

$$\frac{\partial f}{\partial T}=\min_v [(v,\ v)+(c,\ Ac)+(v,\ \operatorname{grad} f)], \tag{22.4}$$

где $\operatorname{grad} f$ есть вектор, компонентами которого служат $\partial f/\partial c_1$, $\partial f/\partial c_2$, ..., $\partial f/\partial c_N$, и где c_1, c_2, ..., c_N — компоненты вектора c.

Легко видеть, что значение v, при котором достигается минимум, равно

$$v=-\operatorname{grad} f/2. \tag{22.5}$$

Это и есть недостающее начальное условие в задаче (22.2).

Чтобы добиться эффективного использования (22.5), отметим, что $f(c,\ T)$ является квадратичной функцией от c, т. е.

$$f(c,\ T)=(c,\ R(T)c). \tag{22.6}$$

Используя это, видим, что

$$v=-R(T)c, \tag{22.7}$$

---

[1]) Последнее из условий (22.2) — это условие трансверсальности. — *Прим. ред.*

[2]) Минимум здесь берется по всем функциям $x(t)$, определенным при $0\leqslant t\leqslant T$ и удовлетворяющим условию $x(0)=c$. — *Прим. ред.*

где $R(T)$ в силу (22.4) удовлетворяет дифференциальному уравнению Риккати[1])

$$R'(T) = A - R^2, \quad R(0) = 0. \tag{22.8}$$

Это уже *задача Коши*, которую можно численно решить самым обычным образом. При возрастании $R(T)$ затруднений не возникает, так как известно, что $R(T)$ стремится к константе при $T \to \infty$.

Подобный результат получается, только с более длинными выкладками, в случае, когда мы стремимся минимизировать по y функционал

$$J(y) = \int_0^T [(x, Ax) + (y, By)]\, dt, \tag{22.9}$$

где

$$\dot{x} = Cx + Dy, \quad x(0) = c. \tag{22.10}$$

23. Инвариантное погружение

Во многих случаях мы встречаемся с системами линейных дифференциальных уравнений вида

$$\begin{aligned} x' &= Ax + By, \quad x(0) = c, \\ y' &= Cx + Dy, \quad y(a) = d. \end{aligned} \tag{23.1}$$

Если эти системы порождаются вариационными задачами типа (22.9), (22.10), то мы знаем, как избежать решения двухточечной краевой задачи. В противном случае мы все же располагаем мощными методами для нахождения недостающего начального значения $y(0)$, минуя решение линейных алгебраических уравнений. Эти методы следуют из теории инвариантного погружения. Они показывают, что если записать

$$y(0) = R(a)c + T(a)d, \tag{23.2}$$

то R и T, как функции от a, также удовлетворяют дифференциальным уравнениям типа Риккати с известными начальными условиями[2]).

---

[1]) Здесь $R(T)$ — симметричная матрица. Уравнение (22.4) превратится в тождество, если подставить в него (22.6), (22.7) и учесть уравнение (22.8). Отметим еще, что начальное условие в (22.8) следует из очевидного условия $f(c, 0) = 0$. — *Прим. ред.*

[2]) См. примечание к разд. 21 гл. II. — *Прим. ред.*

КОММЕНТАРИИ И БИБЛИОГРАФИЯ

1. В области многоточечных задач сделано сравнительно мало. С некоторыми последними результатами и дальнейшими ссылками можно ознакомиться в работе

Вашакмадзе Т. С., О многоточечных линейных краевых задачах, *Сообщ. АН Груз. ССР*, **35**, № 1 (1964), 29—36.

2—4. Мы следуем статье

Kalaba R., On nonlinear differential equations, the maximum operation and monotone convergence, *J. Math. Mech.*, 8 (1959), 519—574.

5. См.

Bellman R., Kagiwada H., Kalaba R., Orbit determination as a multipoint boundary-value problem and quasilinearization, *Proc. Nat. Acad. Sci. USA*, 48 (1962), 1327—1329.

7. Эти результаты содержатся в работах

Bellman R., Successive approximations and computer storage problems in ordinary differential equations, *Comm. ACM*, **4** (1961), 222—223;
Bellman R., Kalaba R., Kotkin B., Some numerical results using quasilinearization for nonlinear two-point boundary-value problems, The RAND Corporation, RM-3113-PR, April, 1962.

9. См.

Bellman R., On the iterative solution of two-point boundary-value problems, *Boll. d'Unione Mate.*, **16** (1961), 145—149;
Bellman R., Brown T. A., On the computational solution of two-point boundary-value problems, *Boll. d'Unione Mate.*, **19** (1964), 121—123.

10. См.

Bellman R., On the computational solution of differential-difference equations, *J. Math. Anal Appl.*, **2** (1961), 108—110.
Bellman R., Kotkin B., On the numerical solution of a differential-difference equation arising in analytic number theory, *Math. Comput.*, **16** (1962), 473—475.

Обсуждение того, каким образом в химиотерапии возникают дифференциально-разностные уравнения, см. в статье

Беллман Р., От химиотерапии к вычислению траекторий, сб. «Математические проблемы в биологии» под редакцией Р. Беллмана, изд-во «Мир», М., 1966,

где можно найти дальнейшие ссылки.

Систематическое обсуждение дифференциально-разностных уравнений см.

Беллман Р., Кук К. Л., Дифференциально-разностные уравнения, изд-во «Мир», М., 1967;
* Мышкис А. Д., Линейные дифференциальные уравнения с запаздывающим аргументом, М.—Л., 1951.

* Пинни Э., Обыкновенные дифференциально-разностные уравнения, ИЛ, М., 1961

12. Мы следуем работам

Bellman R., Cooke K. L., On the computational solution of a class of functional differential equations, *J. Math. Anal. Appl.* (в печати);
Bellman R., Buell J., Kalaba R., Numerical integration of a differential-difference equation with a decreasing time-lag, The RAND Corporation, RM-4375-NIH, December 1964;
Bellman R., Buell J., Kalaba R., Mathematical experimentation in time-lag modulation, The RAND Corporation, RM-4432-NIH, February 1965.

13. Дальнейшее обсуждение дифференциальной аппроксимации содержится в работах

Bellman R., Kalaba R., Kotkin B., Differential approximation applied to the solution of convolution equations, *Math. Comput.*, **18** (1964), 487—491;
Bellman R., Gluss B., Roth R., On the identification of systems and the unscrabling of data: Some problems suggested by neurophysiology, *Proc. Nat. Acad. Sci. USA*, **52** (1964), 1239—1240;
Bellman R., Segmental differential approximation and the «Black box» problem, *J. Math. Anal. Appl.* (в печати);
Кардашов А. А., Анализ качества систем авторегулирования методом понижения порядка дифференциального уравнения, *Автоматика и телемеханика*, **24**, № 8 (1963), 1073—1083.

Дальнейшие применения дифференциальной аппроксимации можно найти в книге

Bellman R., Perturbation techniques in mathematics, physics and engineering, New York, 1964,

и в статьях

Bellman R., Richardson J. M., On some questions arising in the approximate solution of nonlinear differential equations, *Quart. Appl. Math.*, **20** (1963), 333—339;
Bellman R., Renormalization techniques and mean-square averaging—I: Deterministic equations, *Proc. Nat. Acad. Sci. USA*, **47** (1961), 1191—1194;
Bellman R., Perturbation techniques, Международный симпозиум по нелинейным колебаниям, Киев, 1961;
Richardson J. M., Levitt L. C., Systematic linear closure approximation for classical equilibrium statistical mechanics — I: General method, *Bull. Amer. Phys. Soc., Ser.* II, **9** (1964), 225;
Levitt L. C., Richardson J. M., Cohen E. R., Systematic linear closure approximation for classical equilibrium statistical mechanics — II: Application to the calculation of corrections to the Debye-Huckel theory of an electron gas, *Bull. Amer. Phys. Soc., Ser.* II, **9** (1964), 277.

Интересны также работы

Либерман Л. Х., О некоторых задачах теории приближения решений дифференциально-операторных уравнений в гильбертовом пространстве, *Изв. высш. учеб. заведений, сер. мат.*, 3 (1964), 88—92;

Зайденберг Е. Д., О третьем способе статистической линеаризации одного класса нелинейных дифференциальных уравнений, *Автоматика и телемеханика*, **25** (1964), 195—200;
Culver J. W., Mesarovic M. D., Dynamic statistical linearization, AIEE Joint Automatic Control Conference, New York, June 27—29 (1962).

Общее обсуждение восстановления значений функций см. в обзорной статье

Тихомиров В. М., Работы А. Н. Колмогорова по ε-энтропии функциональных классов и суперпозициям функций, *УМН*, **18** № 5 (1963), 55—92,

где можно найти много дальнейших ссылок.

15. Изложение следует работам

Bellman R., Kalaba R., Kotkin B., Differential approximation applied to the solution of convolution equations, *Math. Comput.*, **18** (1964), 487—491;
Bellman R., Kotkin B., A numerical approach to the convolution equations of a mathematical model of chematherapy, The RAND Corporation, RM-3716-NIH, July 1963.

16. Примеры, поясняющие опасность экспоненциальной аппроксимации, см. в книге

Ланцош К., Практические методы прикладного анализа, Справочное руководство, Физматгиз, М., 1961.

18. Этот результат изложен в работе

Bellman R., Glicksberg I., Gross O., On some variational problems occurring in the theory of dynamic programming, *Rend. Circ. Mat. Palermo, Ser.* II (1954), 1—35.

Интересно отметить, что изучение уравнения (18.11), играющего фундаментальную роль в теории марковских процессов решения непрерывного типа, привлекло внимание к этому типу функциональных уравнений. Отсюда возникли результаты работы

Bellman R., Functional equations in the theory of dynamic programming — V: Positivity and quasilinearity, *Proc. Nat. Acad. Sci. USA*, **41** (1955), 743—746,

которые в свою очередь привели к работе

Kalaba R., On nonlinear differential equations, the maximum operation and monotone convergence, *J. Math. Mech.*, **8** (1959), 519—574.

Наконец, изучение уравнений типа (18.11) привело к естественному расширению теории полугрупп, включающему вариационное исчисление, см. замечание в книге

Беллман Р., Динамическое программирование, ИЛ, М., 1960.

19. См.

Polya G., On the mean-value theorem corresponding to a given linear homogeneous differential equation, *Trans. Amer. Math. Soc.*, **24** (1922), 312—324;
Schlesinger J., Handbuch der Theorie der Linearen Differentialgleichungen, **1** (1895), 52, формула (14).

Обобщение на случай дифференциальных уравнений в частных производных, основанное на работах Пойа и других авторов, см.

Hartman P., Wintner A., Mean value theorems and linear operators, *Amer. Math. Monthly*, **62** (1955), 217—221.

Дальнейшие ссылки см. в книге

Беккенбах Э., Беллман Р., Неравенства, изд-во «Мир», М., 1965.

22. Дальнейшее обсуждение этих методов см в книге

Bellman R., Introduction to matrix analysis, New York, 1960 (русский перевод готовится к печати в изд-ве «Наука»),

где можно найти много других ссылок. Авторы настоящей книги, а также Адорно, Беквит, Фришер, Калман, Копске и др. широко использовали эти методы в современной теории управления. См также литературу к разд. 20 гл. II.

23. Обсуждение и применение инвариантного погружения см

Bellman R., Kalaba R., Prestrud M., Invariant imbedding and radiative transfer in slabs of finite thickness, New York, 1963;
Bellman R., Kagiwada H., Kalaba R., Prestrud M., Invariant imbedding and time-dependent transport processes, New York, 1964;
Bellman R., Kalaba R., Wing G. M., Invariant imbedding and mathematical physics — I: Particle processes, *J. Math. Phys.*, **1** (1960), 280—308;
Калаба Р., Инвариантное погружение и анализ процессов, сб. «Общая теория систем» под редакцией Месаровича М., изд-во «Мир», М., 1966;
Wing G. M., An introduction to transport theory, New York, 1962.

О приложении динамического программирования к численному решению плохо обусловленных систем см.

Bellman R., Kalaba R., Lockett J., Dynamic programming and ill-conditioned linear systems, *J. Math. Anal. Appl.*, **10** (1965), 206—215;
Bellman R., Dynamic programming and ill-conditioned linear systems — II, *J. Math. Anal. Appl.* (в печати).

Эти методы оказываются полезными при численном обращении преобразования Лапласа.

Глава V

Дифференциальные уравнения в частных производных

1. Введение

Займемся теперь дифференциальными уравнениями в частных производных. Как и в предыдущих главах, обсудим вначале некоторые вычислительные аспекты, а затем кратко рассмотрим вопросы монотонности приближений и положительности. Эта глава написана на ином уровне, нежели любая из предыдущих или последующих глав, поскольку в ней предполагается знакомство с весьма развитым и трудным искусством решения дифференциальных уравнений в частных производных на цифровых машинах. Читателю, неискушенному в этих кабалистических расчетах, можно посоветовать опустить эту главу.

Нашим первым числовым примером будет нелинейное параболическое дифференциальное уравнение в частных производных, а вторым — нелинейное эллиптическое дифференциальное уравнение в частных производных.

Вычислительная сторона дела здесь гораздо сложнее, чем в случае обыкновенных дифференциальных уравнений. При получении численного решения обыкновенных дифференциальных уравнений с заданными начальными условиями можно использовать методы такой гарантированной точности, которая позволяет быть совершенно уверенным в том, что результаты отражают эффективность используемой методики. Поэтому имеет смысл сравнивать различные методы. При решении дифференциальных уравнений в частных производных оценка фактически осуществляемой численной процедуры представляет собой целую проблему. Влияние ошибок вследствие отбрасывания цифр, ошибок округления, длины машинного слова и фактических вычислительных процедур применяемой электронной машины, влияние различных грубых оценок для производных высших порядков при выборе методов, а также усилия, затрачиваемые при реализации разностных методов высокого порядка, — все это затрудняет получение обоснованных априорных оценок преимуществ разностных методов высокого порядка перед методами низкого порядка, а также сравнение методов различных авторов. Отсюда следует важность проведения определенного числа экспериментов.

2. Параболические уравнения

Начнем с рассмотрения параболических уравнений в частных производных, поскольку мы уже в какой-то степени обсудили их в связи с положительностью линейных дифференциальных операторов второго порядка.

Для задачи определения численного решения дифференциального уравнения в частных производных вида

$$u_t = u_{xx} + g(u, u_x) \tag{2.1}$$

с начальными условиями

$$u(x, 0) = h(x), \quad 0 < x < 1, \tag{2.2}$$

и краевыми условиями

$$u(0, t) = u(1, t) = 0 \tag{2.3}$$

выпишем схему аппроксимации, следуя нашему обычному методу квазилинеаризации:

$$(u_{n+1})_t = (u_{n+1})_{xx} + g(u_n, (u_n)_x) + (u_{n+1} - u_n)\, g_u(u_n, (u_n)_x) + \\ + ((u_{n+1})_x - (u_n)_x)\, g_{u_x}(u_n, (u_n)_x). \tag{2.4}$$

Задаваясь функцией $u_0(x, t)$, получаем для каждого u_n, $n \geqslant 1$, линейное уравнение.

Перейдем теперь непосредственно к вычислительным вопросам.

3. Конкретное уравнение

Рассмотрим уравнение

$$u_t - u_{xx} = (1 + u^2)(1 - 2u), \tag{3.1}$$

которое подобрано так, что оно имеет простое явное решение

$$u(x, t) = \operatorname{tg}(x + t). \tag{3.2}$$

Это решение обладает тем дополнительным преимуществом, что неограниченно возрастает при $x + t \to \pi/2 = 1{,}57 + \ldots$ и может служить для хорошей проверки численных алгоритмов. Обычный итерационный метод Пикара

$$(u_{n+1})_t - (u_{n+1})_{xx} = (1 + u_n^2)(1 - 2u_n) \tag{3.3}$$

сравнивается с квазилинеаризацией в двух треугольниках $0 \leqslant t \leqslant (1 - x)$, $0 \leqslant x \leqslant 1$ и $0 \leqslant t \leqslant 1{,}5 - x$, $0 \leqslant x \leqslant 1{,}5$.

Численные значения получены с использованием стандартных методов разностной аппроксимации, которые будут описаны

ниже. Как упоминалось ранее, использование других вычислительных алгоритмов решения дифференциальных уравнений в частных производных может привести к иной оценке сравниваемых подходов.

4. Разностная аппроксимация

Первый вопрос состоит в выборе разностной аппроксимации членов $u_t - u_{xx}$. Мы использовали приращения $\Delta x = \Delta t = 0{,}01$ и для получения дискретной системы применяли разностный оператор Кранка — Никольсона.

Обозначая через $u_{m,\,n}$ решение соответствующего (3.1) разностного уравнения в точке $(m\Delta x,\ n\Delta t)$ решетки, заменяем (3.1) уравнением

$$\frac{u^{(k+1)}_{m,\,n+1} - u^{(k+1)}_{m,\,n}}{\Delta t} - \frac{1}{2\,(\Delta x)^2}\left\{\left[u^{(k+1)}_{m-1,\,n+1} - 2u^{(k+1)}_{m,\,n+1} + u^{(k+1)}_{m+1,\,n+1}\right] + \right.$$

$$\left. + \left[u^{(k_n)}_{m-1,\,n} - 2u^{(k_n)}_{m,\,n} + u^{(k_n)}_{m+1,\,n}\right]\right\} = \frac{1}{2}\left[f\left(u^{(k)}_{m,\,n+1}\right) + f\left(u^{(k_n)}_{m,n}\right)\right], \quad (4.1)$$

где $f(u) = (1+u^2)\,(1-2u)$, а k_n обозначает окончательное число итераций, необходимое для получения приемлемой аппроксимации величины $u_{m,\,n}$ в узловых точках на прямой $T = n\Delta t$.

В качестве критерия приемлемости аппроксимации был выбран общепринятый критерий: максимальное значение относительного изменения на слое должно быть меньше некоторой предписанной величины; при удовлетворении критерия совершается переход к следующему слою. В данном случае мы требовали, чтобы

$$\max_m \left| \frac{u^{(k)}_{m,\,n} - u^{(k-1)}_{m,\,n}}{u^{(k)}_{m,\,n}} \right| \leqslant 10^{-6}. \quad (4.2)$$

5. Сравнение метода Пикара с квазилинеаризацией

В табл. 5.1 приведено число итераций, которое потребовалось для удовлетворения критерия (4.2). В качестве важного вывода отсюда следует, что в случае треугольника $0 \leqslant t \leqslant 1 - x$, $0 \leqslant x \leqslant 1$ метод Пикара и метод квазилинеаризации потребовали почти одинакового объема вычислений, в то время как в случае, когда интересующая нас область была бóльшим треугольником $0 \leqslant t \leqslant 1{,}5 - x$, $0 \leqslant x \leqslant 1{,}5$, для получения сравнимой точности в случае метода квазилинеаризации понадобилось количество итераций в девять раз меньшее, чем в случае метода Пикара.

Таблица 5.1

Основание треугольника	Метод	Число итераций	t-интервал
1,0	Ньютон	2 1	(0,01, 0,98) (0,98, 0,99)
1,0	Пикар	4 3 2	(0,01, 0,68) (0,68, 0,93) (0,93, 0,99)
1,5	Ньютон	3	(0,01, 0,23)
1,5	Пикар	28	(0,01, 0,50)

6. Градиентные методы

В гл. I мы указали, что один из подходов к решению скалярного уравнения $f(x)=0$ состоит в использовании метода аппроксимации Ньютона — Рафсона. Вместо первоначального статического процесса рассмотрим дискретный динамический процесс, описываемый нелинейным разностным уравнением

$$x_{n+1}-x_n=\frac{-f(x_n)}{f'(x_n)}. \tag{6.1}$$

При соответствующих условиях предел последовательности $\{x_n\}$ существует и является искомым корнем.

Вместо дискретного процесса можно рассмотреть непрерывный процесс, описываемый дифференциальным уравнением

$$\frac{dx}{dt}=f(x) \tag{6.2}$$

или дифференциальным уравнением в форме, более близкой к (6.1):

$$\frac{dx}{dt}=\frac{-f(x)}{f'(x)}. \tag{6.3}$$

Ясно, что существует бесконечно много дифференциальных уравнений, предельная форма которых при $t\to\infty$ совпадает со статическим уравнением $f(x)=0$. Наш выбор будет зависеть от точности, устойчивости, времени и типа расчетов.

Ту же идею можно использовать для изучения общих функциональных уравнений. Рассмотрим двухточечную краевую задачу

$$u''=g(u, u', x), \quad u(0)=u(b)=0. \tag{6.4}$$

Свяжем с этим уравнением процесс, зависящий от времени:

$$\begin{aligned} &u_t = u_{xx} - g(u, u_x, x), \\ &u(0, t) = u(b, t) = 0, \\ &u(x, 0) = w(x), \quad 0 < x < b, \end{aligned} \tag{6.5}$$

где $w(x)$ выбирается довольно произвольно.

Используя квазилинеаризацию, как описано в предыдущих разделах, мы получаем метод иного типа для решения задачи (6.4); при этом может возникнуть (а может и не возникнуть) необходимость решения линейных систем обыкновенных дифференциальных уравнений или линейных систем алгебраических уравнений.

Может оказаться более предпочтительным использование иного подхода к (6.5) в отличие от описанного выше. Например, кажутся весьма эффективными методы полиномиальной аппроксимации.

Отметим далее, что нас совершенно не интересует изменение $u(x, t)$ на всем интервале времени t. Мы интересуемся только «установившимся режимом»

$$u(x) = \lim_{t \to \infty} u(x, t). \tag{6.6}$$

возлагая надежды на то, что он является решением задачи (6.4).

Поскольку мы ожидаем, что это приближение будет экспоненциальным, т. е.

$$u(x, t) \sim u(x) + u_1(x) e^{\lambda_1(x) t} + u_2(x) e^{\lambda_2(x) t} + \ldots, \tag{6.7}$$

мы можем удачно использовать нелинейные методы суммирования для получения отличной экстраполяции для $u(x)$ по $u(x, t)$ при малых t.

Наконец, отметим, что соотношения (6.5) дают лишь один из методов погружения задачи (6.4) в непрерывное семейство процессов. Теория инвариантного погружения позволяет получить и другие подходы.

Однако эти методы в настоящее время увели бы нас далеко в сторону. Мы вернемся к ним в последующих монографиях. Следует особо подчеркнуть, что весьма желательно сочетание нескольких методов и неразумно и ненаучно пытаться исследовать важные проблемы только одним методом. Ссылки на упомянутые подходы читатель найдет в конце главы.

7. Эллиптическое уравнение

Соответствующие исследования были проведены и для эллиптического уравнения

$$u_{xx}+u_{yy}=e^{u}, \tag{7.1}$$

к которому проявляется интерес в магнитной гидродинамике. Граничные условия будут выписаны ниже.

Решение, полученное методом приближений Пикара при помощи рекуррентного соотношения

$$(u_{k+1})_{xx}+(u_{k+1})_{yy}=e^{u_k}, \tag{7.2}$$

мы сравним с решением, полученным методом квазилинеаризации с использованием соотношения

$$(u_{k+1})_{xx}+(u_{k+1})_{yy}-e^{u_k}u_{k+1}=e^{u_k}(1-u_k). \tag{7.3}$$

Проведение аналогичных доказательств монотонности и квадратичной сходимости по образцам, представленным в предыдущих разделах, не вызывает затруднений. Легко также устанавливается требуемая положительность, если воспользоваться вариационным подходом, изложенным в разд. 8 гл. III.

8. Вычислительные аспекты

Мы будем рассматривать прямоугольную область $0\leqslant x\leqslant 1/2$, $0\leqslant y\leqslant 1/4$ с двумя наборами граничных условий: на границе прямоугольника $u=0$ и $u=10$. Как и следует ожидать, в случае $u=10$, где имеют место большие величины градиента, предпочтительнее метод квазилинеаризации.

Мы следуем стандартному подходу. Лапласиан заменяется разностным выражением

$$\frac{u_{m+1,\,n}+u_{m,\,n+1}+u_{m-1,\,n}+u_{m,\,n-1}-4u_{m,\,n}}{h^2}, \tag{8.1}$$

где $u_{m,\,n}$ является дискретным аналогом величины $u(m\Delta x, n\Delta y)$; $h=\Delta x=\Delta y$ берется в данном случае равным 1/64.

Если упорядочить точки $(m,\ n)$ таким образом, что точка $(m,\ n)$ будет предшествовать точке $(m',\ n')$, когда $n<n'$ или когда $n=n'$ и $m<m'$, и обозначить через u_k полученное множество $15\times 31=465$ чисел $\{u_{mn,\,k}\}$, то и метод Пикара и квазилинеаризация приведут к необходимости решения систем большого числа линейных алгебраических уравнений [1]).

[1]) В качестве u_k рассматриваются значения функции u лишь во внутренних точках области (исключая границу), причем k — это номер итерации в соотношениях (7.2), (7.3). — *Прим. ред.*

Нормируя эти уравнения делением всех членов на отрицательный коэффициент при центральном члене $u_{mn,k}$, получаем в каждом случае систему вида

$$(I - L_k - U_k)\, u_{k+1} = b_k. \tag{8.2}$$

Здесь I, L_k и U_k — матрицы размера 465×465; I — единичная матрица, а L_k и U_k — соответственно нижняя и верхняя треугольные матрицы, отвечающие используемому методу разностной аппроксимации. В нашей задаче эти матрицы являются транспонированными друг к другу и содержат только две диагональные линии с ненулевыми элементами; в случае метода Пикара эти элементы тождественно равны 1/4, а в случае метода Ньютона они равны

$$(4 + h^2 \exp u_k)^{-1},$$

где приближение u_k вычислено в центральной точке (m, n) креста, образуемого точками, которые фигурируют в (8.1). Компоненты вектора b_k в (8.2) равны соответствующим значениям, которые принимают правые части равенств (7.2) и (7.3).

Во многих реальных задачах размерность настолько велика, что нельзя воспользоваться методом исключения Гаусса для решения алгебраических систем вида (8.2). Поэтому мы избрали метод последовательной верхней релаксации Янга главным образом из-за того, что он является простейшим итерационным методом, который существенно лучше метода Гаусса — Зейделя, хотя известно, что быстрее сходящиеся методы групповой релаксации и методы переменных направлений ненамного сложнее метода последовательной верхней релаксации.

Применение метода последовательной верхней релаксации непосредственно к системам (8.2) при фиксированном значении k приводит к итеративной формуле

$$(I - \omega L_k)\, u_{r+1,\,k\,1} = \\ = [\omega U_k - (\omega - 1) I]\, u_{r,\,k+1} + \omega b_k, \quad r = 0,\ 1,\ 2,\ \ldots, \tag{8.3}$$

где ω — параметр верхней релаксации, а уравнения для компонент решаются последовательно таким образом, что компоненты вектора $u_{r+1,\,k+1}$ находятся в указанном выше порядке. Однако ясно, что если u_{k+1} не является хорошей аппроксимацией решения уравнения (7.1), то нет смысла прилагать много усилий для получения точного решения системы (8.2). Аналогично, если приближения, полученные из (8.3), слишком грубы из-за слишком раннего обрыва итерационного процесса по r, то нет никаких указаний относительно итерационного процесса по k. Таким образом, на каждом шаге остается открытым вопрос о том,

когда нужно итерировать по k, а когда по r. Мы не предпринимали попыток проведения этого относительно трудного анализа и просто производили итерации по обоим индексам одновременно, используя формулу

$$(I - \omega L_k)\, u_{k,l} = [\omega U_k - (\omega - 1)\, I]\, u_k + \omega b_k, \tag{8.4}$$

причем уравнения для компонент решались в указанном выше порядке. Таким образом, мы видим тесную связь численного метода, используемого при решении уравнений (8.2) или (8.4), со скоростью сходимости их фактического решения; трудности анализа этой ситуации вынуждают нас экспериментировать. Отметим также, что сочетание итерационных методов решения линейных систем (8.2) с итерациями Ньютона или Пикара для решения исходной нелинейной задачи обладает тем преимуществом, что нет необходимости добиваться точного решения системы (8.2) при каждом значении k.

Как и в параболическом случае, критерий для остановки итерационного процесса состоит в том, что максимум абсолютной величины относительного изменения значения функции для двух последовательных итераций не должен превышать 10^{-6}. Если бы эта величина действительно являлась пределом относительной ошибки между численным и точным решениями нелинейного уравнения (7.1), в котором лапласиан заменен выражением (8.1), то этот критерий был бы слишком жестким и точность его выполнения с учетом ошибок вследствие отбрасывания цифр можно было бы снизить до 10^{-4}. Однако эвристический характер правила остановки, когда мы не знаем ни истинного решения, ни действительных границ его, вынуждает нас увеличивать точность выполнения критерия.

В случае итераций Пикара, когда L_k и U_k не зависят от k, если бы мы пренебрегли зависимостью b_k от k, то получили бы, что наибыстрейшая сходимость достигается при теоретически оптимальном значении ω, равном 1,732. Для итерации Ньютона оптимальное значение ω было бы немного меньшим, если пренебречь изменением L_k, U_k и b_k от итерации к итерации, что мы, несомненно, не имеем права делать. Однако в случае нулевых краевых условий для метода Пикара, где L_k и U_k постоянны, наилучшее значение (или значения) ω, несомненно, близко к указанному выше теоретическому значению, так как u_k во всей области меняется очень слабо (значение u в центре области, полученное в окончательном решении, равно 0,007071 с точностью до четырех значащих цифр). Поэтому в данном случае мы проводили эксперимент только со значениями $\omega = 1{,}70$ и 1,74. В случаях применения метода Пикара, когда $u = 10$ на границе, такой уверенности не было из-за больших изменений u_k в

заданной области, несомненно, приводящих к большим изменениям от итерации к итерации для каждой точки области. Разумеется, в случае метода Ньютона имеется еще меньше оснований рассчитывать на оптимальное значение ω, так как L_k и U_k изменяются от итерации к итерации. Поэтому в этих случаях для оценки оптимального значения ω был приведен ряд экспериментов с различными значениями ω. В табл. 5.2 приведены полученные результаты.

Таблица 5.2

Значение ω	Число итераций	Значение $u^{(k)}$ в точке $\left(\frac{1}{4}, \frac{1}{8}\right)$
	$u \equiv 0$ (метод Пикара)	
1,74	64	—0,007071
1,70	77	—0,007071
	$u \equiv 0$ (метод Ньютона)	
1,95	258	—0,007071
1,74	64	—0,007071
1,70	77	—0,007071
1,50	154	—0,007071
1,00	426	—0,007071
	$u \equiv 10$ (метод Пикара)	
1,55	63	5,65995
1,50	53	5,65995
1,40	66	5,65994
1,35	71	5,65993
1,30	76	5,65993
1,20	78	5,65992
1,10	106	5,65997
1,00	134	5,65998
	$u \equiv 10$ (метод Ньютона)	
1,74	58	5,65995
1,70	51	5,65995
1,65	46	5,65995
1,60	42	5,65995
1,55	41	5,65995
1,50	50	5,65995
1,40	66	5,65996
1,00	148	5,65998

Из таблицы видно, что в случае нулевых граничных условий метод Ньютона, обладающий свойством квадратичной сходимости, не дал никаких преимуществ по сравнению с методом Пикара, сходящимся, как геометрическая прогрессия. В самом деле, одинаковые значения ω дали в обоих случаях наибыстрей-

шую сходимость. С другой стороны, когда $u=10$ на границе, что влечет за собой резкое изменение u вблизи границы, метод Ньютона оказался более выгодным, так как для решения потребовалось всего 41 итерация, в то время как для решения методом Пикара потребовалось 53 итерации. Отметим еще любопытный факт: для метода Пикара оптимальное значение ω получается несколько меньшим.

Для сравнения с последовательной верхней релаксацией мы включили в рассмотрение случай релаксации Гаусса — Зейделя, получающийся при $\omega=1$. Как и ожидалось, метод последовательной верхней релаксации сходится примерно в 2,5—7 раз быстрее, чем метод релаксации Гаусса — Зейделя.

Заметим попутно, что точки, в которых достигается максимум относительного изменения u_h, неизменно находились в непосредственной близости от начала координат. В большинстве случаев это имело место в точке (1/64, 1/64).

9. Положительность и монотонность

Любое детальное обсуждение свойств положительности и монотонности дифференциального оператора в частных производных увело бы нас далеко вглубь теории дифференциальных уравнений в частных производных. Поэтому остановимся лишь на нескольких простых свойствах методов, которые мы уже применяли.

С помощью метода конечных разностей, который мы использовали в разд. 6 гл. III, легко устанавливается желаемое свойство неотрицательности решения уравнения

$$\begin{aligned} u_t &= u_{xx} + f(x, t), && t > 0, \\ u(x, 0) &= g(x), && 0 \leqslant x \leqslant 1, \\ u(0, t) &= u(1, t) = 0, && t > 0, \end{aligned} \tag{9.1}$$

где $f,\ g \geqslant 0$, и решений подобных уравнений более высокой размерности.

Для получения соответствующего результата для эллиптического уравнения в частных производных вида

$$u_{xx} + u_{yy} + q(x, y)\, u = f(x, y) \tag{9.2}$$

можно либо использовать параболическое дифференциальное уравнение в частных производных, как в разд. 6 гл. III, либо исходить из вариационной задачи минимизации функционала

$$\int \left[u_x^2 + u_y^2 - q(x, y)\, u^2 - 2f(x, y)\, u\right] dA. \tag{9.3}$$

Таким способом можно получить не только неотрицательность функции Грина, но также и аналог свойства осцилляции.

10. Уравнение Хопфа — Лакса

Рассмотрим уравнение сохранения

$$u_t + f(u)_x = 0 \tag{10.1}$$

с начальными условиями $u(x, 0) = g(x)$, $T \geqslant t \geqslant 0$. Особенно интересен случай $f(u) = u^2/2$, так как тогда уравнение является предельным случаем уравнения Бейтмана — Бюргерса

$$u_t + uu_x = \varepsilon u_{xx}, \tag{10.2}$$

играющего важную роль в некоторых упрощенных математических моделях турбулентности.

Пусть $f(u)$ строго выпукла. Запишем (10.1) в виде

$$u_t + f'(u)\, u_x = 0 \tag{10.3}$$

и введем функцию $U(x, t)$ с помощью соотношения $U_x = u$. Тогда уравнение (10.3) примет вид

$$U_{xt} + f'(U_x)\, U_{xx} = 0. \tag{10.4}$$

Интегрируя (10.4) по x при заданных начальных условиях, получаем

$$U_t + f(U_x) = 0. \tag{10.5}$$

Соответствующее линейное уравнение имеет вид

$$w_t + f(v) + (w_x - v) f'(v) = 0, \quad w(x, 0) = G(x). \tag{10.6}$$

Используя легко проверяемое свойство положительности оператора

$$\frac{\partial}{\partial t} + f'(v)\, \frac{\partial}{\partial x},$$

находим явное решение уравнения (10.5) в виде

$$U = \min_v w(x, t, v), \tag{10.7}$$

где w — решение уравнения (10.6).

Полагая v константой, получаем простое явное представление решения, найденное ранее Лаксом в общем случае и Хопфом в частном случае $f(u) = u^2/2$. В нашем случае уравнение можно свести к линейному уравнению теплопроводности и решить в явном виде.

КОММЕНТАРИИ И БИБЛИОГРАФИЯ

1. Численные результаты были получены совместно с М. Янкозой. См.

Bellman R., Juncosa M., Kalaba R., Some numerical experiments using Newton's method for nonlinear parabolic and elliptic boundary-value problems, *Comm. ACM*, **4** (1961), 187—191,

где приводятся дальнейшие ссылки.

6. Детальное обсуждение градиентных методов см.

Rosenbloom P., The method of steepest descent, *Proc. Symp. Appl. Math.*, **6** (1956), 127—176.

Краткий обзор применения нелинейных методов суммируемости см.

Bellman R., Kalaba R., A note on nonlinear summability techniques in invariant imbedding, *J. Math. Anal. Appl.*, **6** (1963), 465—472.

Что касается использования полиномиальной аппроксимации при численном решении дифференциальных уравнений в частных производных, см.

Bellman R., Kalaba R., Kotkin B., On a new approach to the computational solution of partial differential equations, *Proc. Nat. Acad. Sci. USA*, **48** (1962), 1325—1327;
Bellman R., Some questions concerning difference approximations to partial differential equations, *Boll. d'Unione Mate.*, **17** (1962), 188—190.

9. Ряд ссылок на свойства положительности решений дифференциальных уравнений в частных производных содержится в книге

Беккенбах Э., Беллман Р., Неравенства, изд-во «Мир», М., 1965.

Эти свойства в свою очередь тесно связаны с «принципами максимума», и любое их детальное обсуждение увело бы нас вглубь современной теории дифференциальных уравнений в частных производных.

10. См.

Lax P., Hyperbolic systems of conservation laws, II, *Comm. Pure Appl. Math.*, **10** (1957), 537—566;
Kalaba R., On nonlinear differential equations, the maximum operation and monotone convergence, *J. Math. Mech.*, **8** (1959), 519—574;
Bellman R., Richardson J. M., Azen S., On new and direct computational approaches to some mathematical models of turbulence, *Quart. Appl. Math.* (в печати).

Чтобы избежать слишком обширных рассмотрений, мы не обсуждали связь задачи минимизации функционалов вида

$$J(u) = \int_0^T h(u, u')dt$$

с дифференциальными уравнениями в частных производных вида $u_t = g(u, u_x t)$. См.

Conway E. D., Hopf E., Hamilton's theory and generalized solutions of the Hamilton-Jacobi equation, *J. Math. Mech.*, **13** (1964), 939—986.

Это в свою очередь приводит нас к приближению в пространстве политик в вариационном исчислении.

О приложении динамического программирования к изучению параболических дифференциальных уравнений в частных производных см.

Fleming W. H., The Cauchy problem for degenerate parabolic equations, *J. Math. Mech.*, **13** (1964), 987—1008.

См. также гл. 4 книги

Беллман Р., Процессы регулирования с адаптацией, изд-во «Наука», М., 1964.

* Отметим, что решение задачи Коши для уравнения (10.1), а также для более общего уравнения

$$u_t + f(t, x, u)_x = 0$$

при условии $f_{uu} \geqslant 0$ было получено в явном виде в работе

Олейник О. А., О задаче Коши для нелинейных уравнений в классе разрывных функций, *ДАН СССР*, **95**, № 3 (1954), 451—455;

подробное изложение этих результатов содержится также в работе

Олейник О. А., Разрывные решения нелинейных дифференциальных уравнений, *УМН*, **12**, № 3 (1957), 3—73.

По поводу этих уравнений см. также

Курант Р., Уравнения с частными производными, изд-во «Мир», 1964, стр. 153—158.

Глава VI

Приложения к физике, технике и биологии

1. Введение

В этой главе мы хотим кратко описать несколько типичных приложений квазилинеаризации к характерным задачам физики, техники и биологии. Мы рассмотрим по очереди задачи оптимального проектирования и управления, обратные задачи переноса излучения, вопросы, связанные с определением орбит, проблемы идентификации систем и управления с адаптацией, анализ периодических зависимостей и, наконец, обратную задачу кардиологии.

2. Оптимальное проектирование и управление

Основная задача современной теории управления заключается в минимизации скалярного функционала

$$J(y) = \int_0^T g(x,\ y)\,dt \tag{2.1}$$

на множестве M-мерных вектор-функций $y(t)$, где N-мерный вектор x и вектор y связаны дифференциальным уравнением

$$\frac{dx}{dt} = h(x,\ y), \quad x(0) = c. \tag{2.2}$$

Часто ставится более общая задача, состоящая в минимизации скалярного функционала

$$J(y,\ a) = \int_0^T g(x,\ y,\ a)\,dt + \varphi(a) \tag{2.3}$$

на множестве M-мерных вектор-функций $y(t)$ и K-мерных векторов a, где x, y и a связаны дифференциальным уравнением

$$\frac{dx}{dt} = h(x,\ y,\ a), \quad x(0) = c. \tag{2.4}$$

В такой постановке можно сформулировать проблему проектирования, поскольку a можно рассматривать как вектор конструктивных параметров. Мы не будем здесь обсуждать

более общие задачи, включающие функциональные дифференциальные уравнения, а также уравнения в частных производных.

При рассмотрении данной вариационной задачи с привлечением квазилинеаризации можно действовать, как мы знаем, двумя различными способами. Мы можем, отправляясь от начального приближения y_0, a_0, получить x_0 с помощью (2.4):

$$\frac{dx_0}{dt} = h(x_0, y_0, a), \quad x_0(0) = c. \tag{2.5}$$

Чтобы получить x_1, y_1, a_1, мы рассматриваем уравнение

$$\frac{dx_1}{dt} = h_1(x_1, y_1, a_1; x_0, y_0, a_0), \quad x_1(0) = c, \tag{2.6}$$

где через h_1 обозначено разложение в окрестности точки (x_0, y_0, a_0) функции $h(x, y, a)$, включающее только *линейные* члены, и ставим задачу минимизации на множестве векторов y_1 и a_1 функционала

$$\int_0^T g_2(x_1, y_1, a_1, x_0, y_0, a_0)\, dt, \tag{2.7}$$

где через g_2 обозначено разложение в окрестности точки (x_0, y_0, a_0) функции $g(x, y, a)$ вплоть до квадратичных членов.

Легко видеть, что вариационные уравнения (уравнения Эйлера) будут линейны относительно x_1, y_1 и a_1.

С другой стороны, мы можем использовать стандартную вариационную процедуру для получения уравнения Эйлера и затем применить квазилинеаризацию к этому уравнению. Для того чтобы унифицировать рассмотрение, мы будем трактовать постоянный вектор a как решение дифференциального уравнения

$$\frac{da}{dt} = 0 \tag{2.8}$$

с неизвестным начальным условием.

Затем, вводя множители Лагранжа $\lambda(t)$ и $\mu(t)$, мы рассматриваем новый функционал

$$\int_0^T [g(x, y, a) + (\lambda(t), \dot{x} - h(x, y, a)) + (\mu(t), \dot{a})]\, dt + \varphi(a(0)) \tag{2.9}$$

и стандартным образом выписываем соответствующие дифференциальные уравнения и граничные условия[1]).

[1]) Эти уравнения и граничные условия получаем, приравнивая нулю первую вариацию функционала (2.9). См., например, Гельфанд И. М. и Фомин С. В., Вариационное исчисление, Физматгиз, М., 1961. — *Прим. ред.*

3. Пример

Рассмотрим для примера задачу минимизации функционала

$$J(y,\ a)=\frac{1}{2}\int_0^1 (x^2+y^2)\,dt+\frac{a^2}{2} \tag{3.1}$$

на множестве функций y и параметра a, где

$$\frac{dx}{dt}=-ax+y,\quad x(0)=c,\quad \frac{da}{dt}=0. \tag{3.2}$$

Уравнения Эйлера и граничные условия имеют вид

$$\begin{aligned}
&\dot{x}=-ax+y, && x(0)=c,\\
&\dot{\lambda}=x+\lambda a, && \lambda(1)=0,\\
&\dot{a}=0, && \mu(0)=a(0),\\
&\dot{\mu}=\lambda x, && \mu(1)=0,\\
&y=\lambda.
\end{aligned} \tag{3.3}$$

Соотношение $y=\lambda$ позволяет исключить λ из рассмотрения. В качестве начального приближения возьмем

$$x_0(t)=1,\quad y_0(t)=0,\quad \mu_0(t)=0,\quad a_0=0. \tag{3.4}$$

Тогда любой из двух методов, рассмотренных в разд. 2, приводит к следующей схеме последовательных приближений:

$$\begin{aligned}
&\dot{x}_{n+1}=-a_n x_{n+1}+y_{n+1}-x_n(a_{n+1}-a_n),\\
&\qquad x_{n+1}(0)=c,\\
&\dot{y}_{n+1}=x_{n+1}+a_n y_{n+1}+y_n(a_{n+1}-a_n),\\
&\qquad y_{n+1}(1)=0,\\
&\dot{\mu}_{n+1}=x_{n+1}y_n+x_n(y_{n+1}-y_n),\\
&\qquad \mu_{n+1}(0)=a_{n+1}(0),\\
&\dot{a}_{n+1}=0,\quad \mu_{n+1}(1)=0.
\end{aligned} \tag{3.5}$$

Применение рассмотренных выше стандартных методов численного определения последовательности $\{x_n,\ y_n,\ a_n\}$ не вызывает затруднений. В следующем разделе мы приведем некоторые характерные примеры вычислений.

Таблица 6.1

Некоторые численные результаты

Итерация 1

T	X	Y	MU	A
0,	$0{,}100000 \cdot 10^{1}$	$-0{,}662196 \cdot 10^{0}$	$0{,}283436 \cdot 10^{0}$	$0{,}283436 \cdot 10^{0}$
0,1000	$0{,}910758 \cdot 10^{0}$	$-0{,}566718 \cdot 10^{0}$	$0{,}223168 \cdot 10^{0}$	$0{,}283436 \cdot 10^{0}$
0,2000	$0{,}830158 \cdot 10^{0}$	$-0{,}479745 \cdot 10^{0}$	$0{,}170912 \cdot 10^{0}$	$0{,}283436 \cdot 10^{0}$
0,3000	$0{,}757867 \cdot 10^{0}$	$-0{,}400410 \cdot 10^{0}$	$0{,}126965 \cdot 10^{0}$	$0{,}283436 \cdot 10^{0}$
0,4000	$0{,}693161 \cdot 10^{0}$	$-0{,}327919 \cdot 10^{0}$	$0{,}906022 \cdot 10^{-1}$	$0{,}283436 \cdot 10^{0}$
0,5000	$0{,}635393 \cdot 10^{0}$	$-0{,}261546 \cdot 10^{0}$	$0{,}611771 \cdot 10^{-1}$	$0{,}283436 \cdot 10^{0}$
0,6000	$0{,}583983 \cdot 10^{0}$	$-0{,}200628 \cdot 10^{0}$	$0{,}381112 \cdot 10^{-1}$	$0{,}283436 \cdot 10^{0}$
0,7000	$0{,}538418 \cdot 10^{0}$	$-0{,}144555 \cdot 10^{0}$	$0{,}208900 \cdot 10^{-1}$	$0{,}283436 \cdot 10^{0}$
0,8000	$0{,}498242 \cdot 10^{0}$	$-0{,}927650 \cdot 10^{-1}$	$0{,}905750 \cdot 10^{-2}$	$0{,}283436 \cdot 10^{0}$
0,9000	$0{,}463053 \cdot 10^{0}$	$-0{,}447402 \cdot 10^{-1}$	$0{,}221157 \cdot 10^{-2}$	$0{,}283436 \cdot 10^{0}$
1,0000	$0{,}432498 \cdot 10^{0}$	$0{,}167638 \cdot 10^{-7}$	$0{,}125729 \cdot 10^{-7}$	$0{,}283436 \cdot 10^{0}$

Итерация 2

T	X	Y	MU	A
0,	$0{,}100000 \cdot 10^{1}$	$-0{,}640197 \cdot 10^{0}$	$0{,}252000 \cdot 10^{0}$	$0{,}232000 \cdot 10^{0}$
0,1000	$0{,}917906 \cdot 10^{0}$	$-0{,}557950 \cdot 10^{0}$	$0{,}174601 \cdot 10^{0}$	$0{,}232000 \cdot 10^{0}$
0,2000	$0{,}845465 \cdot 10^{0}$	$-0{,}481894 \cdot 10^{0}$	$0{,}128792 \cdot 10^{0}$	$0{,}232000 \cdot 10^{0}$
0,3000	$0{,}781899 \cdot 10^{0}$	$-0{,}410979 \cdot 10^{0}$	$0{,}925033 \cdot 10^{-1}$	$0{,}232000 \cdot 10^{0}$
0,4000	$0{,}726531 \cdot 10^{0}$	$-0{,}344449 \cdot 10^{0}$	$0{,}640778 \cdot 10^{-1}$	$0{,}232000 \cdot 10^{0}$
0,5000	$0{,}678772 \cdot 10^{0}$	$-0{,}281596 \cdot 10^{0}$	$0{,}421680 \cdot 10^{-1}$	$0{,}232000 \cdot 10^{0}$
0,6000	$0{,}638114 \cdot 10^{0}$	$-0{,}221749 \cdot 10^{0}$	$0{,}257069 \cdot 10^{-1}$	$0{,}232000 \cdot 10^{0}$
0,7000	$0{,}604125 \cdot 10^{0}$	$-0{,}164271 \cdot 10^{0}$	$0{,}138468 \cdot 10^{-1}$	$0{,}232000 \cdot 10^{0}$
0,8000	$0{,}576444 \cdot 10^{0}$	$-0{,}108548 \cdot 10^{0}$	$0{,}592470 \cdot 10^{-2}$	$0{,}232000 \cdot 10^{0}$
0,9000	$0{,}554777 \cdot 10^{0}$	$-0{,}539856 \cdot 10^{-1}$	$0{,}143373 \cdot 10^{-2}$	$0{,}232000 \cdot 10^{0}$
1,0000	$0{,}538896 \cdot 10^{0}$	$-0{,}223517 \cdot 10^{-7}$	$0{,}814907 \cdot 10^{-8}$	$0{,}232000 \cdot 10^{0}$

Итерация 3

T	X	Y	MU	A
0,	$0{,}100000 \cdot 10^{1}$	$-0{,}640414 \cdot 10^{0}$	$0{,}232393 \cdot 10^{0}$	$0{,}232393 \cdot 10^{0}$
0..000	$0{,}917788 \cdot 10^{0}$	$-0{,}558586 \cdot 10^{0}$	$0{,}175557 \cdot 10^{0}$	$0{,}232393 \cdot 10^{0}$
0,2000	$0{,}845306 \cdot 10^{0}$	$-0{,}482596 \cdot 10^{0}$	$0{,}129699 \cdot 10^{0}$	$0{,}232393 \cdot 10^{0}$
0,3000	$0{,}781742 \cdot 10^{0}$	$-0{,}411697 \cdot 10^{0}$	$0{,}933488 \cdot 10^{-1}$	$0{,}232393 \cdot 10^{0}$
0,4000	$0{,}726424 \cdot 10^{0}$	$-0{,}345141 \cdot 10^{0}$	$0{,}648322 \cdot 10^{-1}$	$0{,}232393 \cdot 10^{0}$
0,5000	$0{,}678770 \cdot 10^{0}$	$-0{,}282227 \cdot 10^{0}$	$0{,}428066 \cdot 10^{-1}$	$0{,}232393 \cdot 10^{0}$
0,6000	$0{,}638276 \cdot 10^{0}$	$-0{,}222290 \cdot 10^{0}$	$0{,}262037 \cdot 10^{-1}$	$0{,}232393 \cdot 10^{0}$
0,7000	$0{,}604515 \cdot 10^{0}$	$-0{,}164697 \cdot 10^{0}$	$0{,}141850 \cdot 10^{-1}$	$0{,}232393 \cdot 10^{0}$
0,8000	$0{,}577132 \cdot 10^{0}$	$-0{,}108842 \cdot 10^{0}$	$0{,}610581 \cdot 10^{-2}$	$0{,}232393 \cdot 10^{0}$
0,9000	$0{,}555837 \cdot 10^{0}$	$-0{,}541356 \cdot 10^{-1}$	$0{,}148800 \cdot 10^{-2}$	$0{,}232393 \cdot 10^{0}$
1,0000	$0{,}540406 \cdot 10^{0}$	$0{,}558794 \cdot 10^{-8}$	$0{,}325963 \cdot 10^{-8}$	$0{,}232393 \cdot 10^{0}$

Итерация 4

T	X	Y	MU	A
0,	$0{,}100000 \cdot 10^{1}$	$-0{,}640448 \cdot 10^{0}$	$0{,}232504 \cdot 10^{0}$	$0{,}232504 \cdot 10^{0}$
0,1000	$0{,}917827 \cdot 10^{0}$	$-0{,}558566 \cdot 10^{0}$	$0{,}175554 \cdot 10^{0}$	$0{,}232504 \cdot 10^{0}$
0,2000	$0{,}845336 \cdot 10^{0}$	$-0{,}482579 \cdot 10^{0}$	$0{,}129696 \cdot 10^{0}$	$0{,}232504 \cdot 10^{0}$
0,3000	$0{,}781763 \cdot 10^{0}$	$-0{,}411682 \cdot 10^{0}$	$0{,}933461 \cdot 10^{-1}$	$0{,}232504 \cdot 10^{0}$
0,4000	$0{,}726438 \cdot 10^{0}$	$-0{,}345128 \cdot 10^{0}$	$0{,}648299 \cdot 10^{-1}$	$0{,}232504 \cdot 10^{0}$
0,5000	$0{,}678777 \cdot 10^{0}$	$-0{,}288216 \cdot 10^{0}$	$0{,}428047 \cdot 10^{-1}$	$0{,}232504 \cdot 10^{0}$
0,6000	$0{,}638277 \cdot 10^{0}$	$-0{,}222281 \cdot 10^{0}$	$0{,}262024 \cdot 10^{-1}$	$0{,}232504 \cdot 10^{0}$
0,7000	$0{,}604510 \cdot 10^{0}$	$-0{,}164690 \cdot 10^{0}$	$0{,}141842 \cdot 10^{-1}$	$0{,}232504 \cdot 10^{0}$
0,8000	$0{,}577121 \cdot 10^{0}$	$-0{,}108838 \cdot 10^{0}$	$0{,}610543 \cdot 10^{-2}$	$0{,}232504 \cdot 10^{0}$
0,9000	$0{,}555821 \cdot 10^{0}$	$-0{,}541333 \cdot 10^{-1}$	$0{,}148790 \cdot 10^{-2}$	$0{,}232504 \cdot 10^{0}$
1,0000	$0{,}540384 \cdot 10^{0}$	$0{,}335276 \cdot 10^{-7}$	$0{,}791624 \cdot 10^{-8}$	$0{,}232504 \cdot 10^{0}$

Итерация 5

T	X	Y	MU	A
0,	$0{,}100000 \cdot 10^{1}$	$-0{,}640446 \cdot 10^{0}$	$0{,}232506 \cdot 10^{0}$	$0{,}232506 \cdot 10^{0}$
0,1000	$0{,}917826 \cdot 10^{0}$	$-0{,}558566 \cdot 10^{0}$	$0{,}175554 \cdot 10^{0}$	$0{,}232506 \cdot 10^{0}$
0,2000	$0{,}845336 \cdot 10^{0}$	$-0{,}482578 \cdot 10^{0}$	$0{,}129696 \cdot 10^{0}$	$0{,}232506 \cdot 10^{0}$
0,3000	$0{,}781763 \cdot 10^{0}$	$-0{,}411681 \cdot 10^{0}$	$0{,}933459 \cdot 10^{-1}$	$0{,}232506 \cdot 10^{0}$
0,4000	$0{,}726438 \cdot 10^{0}$	$-0{,}345128 \cdot 10^{0}$	$0{,}648297 \cdot 10^{-1}$	$0{,}232506 \cdot 10^{0}$
0,5000	$0{,}678777 \cdot 10^{0}$	$-0{,}282215 \cdot 10^{0}$	$0{,}428046 \cdot 10^{-1}$	$0{,}232506 \cdot 10^{0}$
0,6000	$0{,}638276 \cdot 10^{0}$	$-0{,}222280 \cdot 10^{0}$	$0{,}262024 \cdot 10^{-1}$	$0{,}232506 \cdot 10^{0}$
0,7000	$0{,}604510 \cdot 10^{0}$	$-0{,}164690 \cdot 10^{0}$	$0{,}141842 \cdot 10^{-1}$	$0{,}232506 \cdot 10^{0}$
0,8000	$0{,}577121 \cdot 10^{0}$	$-0{,}108838 \cdot 10^{0}$	$0{,}610542 \cdot 10^{-2}$	$0{,}232506 \cdot 10^{0}$
0,9000	$0{,}555820 \cdot 10^{0}$	$-0{,}541333 \cdot 10^{-1}$	$0{,}148790 \cdot 10^{-2}$	$0{,}232506 \cdot 10^{0}$
1,0000	$0{,}540383 \cdot 10^{0}$	$0{,}428408 \cdot 10^{-7}$	$0{,}162981 \cdot 10^{-7}$	$0{,}232506 \cdot 10^{0}$

4. Численные результаты

Полагая $c=1$ и задавая x_0, y_0, μ_0 и a_0 согласно (3.4), мы получаем результаты, приведенные в табл. 6.1. Мы видим, что оптимальное значение a равно

$$a \cong 0{,}2325 \tag{4.1}$$

(программа приведена в приложении 2).

Чтобы проверить этот результат, положим $a=0{,}2325$ и решим вариационную задачу, описываемую соотношениями (3.1) и (3.2). Это сделать легко вследствие простоты ее аналитической структуры; таким же способом мы получаем результаты для ряда близких значений a (рис. 6.1). Из рисунка видно, что значение a, равное 0,2325, на самом деле оптимально.

5. Обратная задача переноса излучения

Теперь рассмотрим один из вариантов так называемых «обратных» задач. В двух предыдущих книгах из этой серии

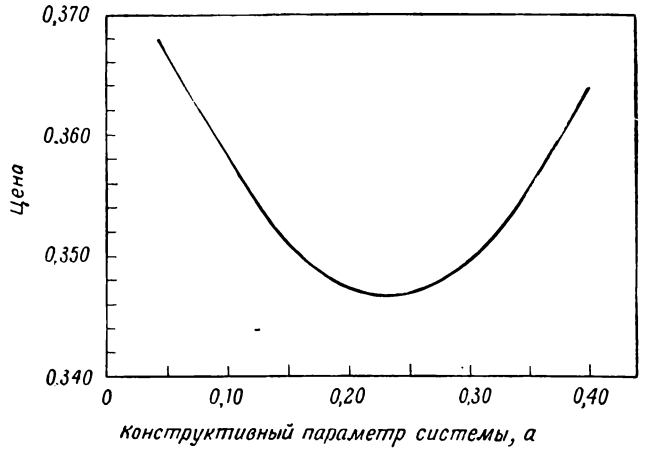


Рис. 6.1. Цена как функция конструктивного параметра.

(см. библиографию к разд. 5 гл. VI) мы обсуждали задачу определения отраженного потока от неоднородной плоско-параллельной неизлучающей и изотропно рассеивающей атмосферы конечной оптической толщины τ_1, причем оптические свойства атмосферы зависят только от оптической высоты τ (рис. 6.2).

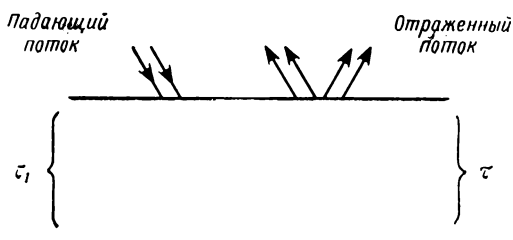


Рис. 6.2. Физическая схема.

Пусть параллельные лучи света, плотность потока которого на единицу площади, перпендикулярной направлению их распространения, равна π, падают на поверхность в направлении, характеризуемом величиной μ_0, где μ_0, как обычно, есть косинус угла, образованного внутренней нормалью с поверхностью. Для

простоты предположим, что нижняя поверхность полностью поглощает свет.

Сформулируем этот процесс в терминах инвариантного погружения. Обозначим через $r(\mu, \mu_0, \tau_1)$ интенсивность диффузионно отраженного света в направлении arc cos μ и положим $R(\mu, \mu_0, \tau_1) = 4\mu r$. Тогда R удовлетворяет интегро-дифференциальному уравнению

$$\frac{\partial R}{\partial \tau_1} = -(\mu^{-1} + \mu_0^{-1}) R + \\ + \lambda(\tau_1)\left[1 + \frac{1}{2}\int_0^1 R(\mu, \mu', \tau_1)\frac{d\mu'}{\mu'}\right]\left[1 + \frac{1}{2}\int_0^1 R(\mu', \mu_0, \tau_1)\frac{d\mu'}{\mu'}\right] \quad (5.1)$$

с начальным условием $R(\mu, \mu_0, 0) = 0$. Функция $\lambda(\tau)$ есть альбедо для однократного рассеивания.

Чтобы получить численное решение уравнения (5.1) на цифровой вычислительной машине, применим гауссову квадратуру порядка N к интегралам и рассмотрим образующуюся в результате систему обыкновенных дифференциальных уравнений [1])

$$\frac{d}{d\tau_1} R_{ij}(\tau_1) = -(\mu_i^{-1} + \mu_j^{-1}) R_{ij} + \\ + \lambda(\tau_1)\left[1 + \frac{1}{2}\sum_{k=1}^{N} R_{ik}(\tau_1)\frac{w_k}{\mu_k}\right]\left[1 + \frac{1}{2}\sum_{k=1}^{N} R_{kj}(\tau_1)\frac{w_k}{\mu_k}\right], \quad R_{ij}(0) = 0, \quad (5.2)$$

где

$$R_{ij}(\tau_1) \equiv R(\mu_i, \mu_j, \tau_1). \quad (5.3)$$

Предположим, с другой стороны, что заданы значения $R_{ij}(\tau_1)$, $i, j = 1, 2, \ldots, N$, и требуется определить функцию $\lambda(\tau)$, $0 \leqslant \tau \leqslant \tau_1$. Вопрос такого типа естественным образом возникает в астрофизике в связи с различными экспериментальными исследованиями.

6. Аналитическая формулировка

Чтобы приняться за решение задачи подобного типа, мы должны задать некоторые аналитические свойства $\lambda(\tau)$, иными словами, мы должны увеличить объем информации о природе физического процесса. Для иллюстрации рассмотрим случай, когда среда состоит по существу из двух слоев с разными постоянными величинами альбедо, причем слои разделены тонкой

[1]) В формуле (5.2) числа w_k — это коэффициенты квадратурной формулы Гаусса. См., например, Березин И. С. и Жидков Н. П., Методы вычислений, т. I, Физматгиз, М., 1962. — *Прим. ред.*

зоной, в которой осуществляется быстрый переход от одного значения альбедо к другому[1]). Пусть альбедо имеет вид

$$\lambda(\tau) = a + b \operatorname{th}[10(\tau - c)], \tag{6.1}$$

где a, b и c — постоянные, которые должны быть определены. Мы видим, что $\lambda_1 \cong a - b$ в слое 1 и $\lambda_2 \cong a + b$ в слое 2.

По известным N^2 значениям $b_{ij} = r_{ij}(\tau_1)$ мы хотели бы определить три параметра a, b и c и толщину τ_1.

Для этого мы ищем значения a, b, c и τ_1, минимизирующие среднеквадратичную ошибку

$$S = \sum_{i,\,j} [r_{ij}(\tau_1) - b_{ij}]^2, \tag{6.2}$$

где $r_{ij}(\tau_1) = R_{ij}(\tau_1)/4\mu_i$ и $R_{ij}(\tau_1)$ удовлетворяет уравнению (5.2).

Эта вариационная задача рассматривается ниже с помощью методов квазилинеаризации, которые мы уже ранее обсуждали.

7. Численные результаты

Начнем с проведения «наблюдений». Для этого выбираем $a = 0{,}5$, $b = 0{,}1$, $c = 0{,}5$ и интегрируем до толщины $\tau_1 = 1{,}0$, причем $N = 7$. Теперь осуществляем три вида численных экспериментов:

а) При заданных a, b и τ_1 определяем c — высоту внутреннего слоя раздела.

б) Определяем полную толщину τ_1 при заданных a, b и c.

в) Определяем a, b и c при заданном τ_1.

Как видно из краткой сводки результатов, в случаях (а) и (в) (табл. 6.2 и 6.3) метод квазилинеаризации работает хо-

Таблица 6.2

Последовательные приближения к высоте внутреннего слоя c

Приближение	Серия 1	Серия 2	Серия 3
0	0,2	0,8	0,0
1	0,62	0,57	
2	0,5187	0,5024	не сходится
3	0,500089	0,499970	
4	0,499990	0,499991	
Истинное значение	0,5	0,5	0,5

[1]) На важность таких задач внимание авторов обратил С. Уэно из университета в Киото, Япония.

Таблица 6.3

Последовательные приближения к λ_1, λ_2 и c.

Приближение	$\lambda_1 = a - b$	$\lambda_2 = a + b$	c
0	0,51	0,69	0,4
1	0,4200	0,6052	0,5038
2	0,399929	0,599995	0,499602
3	0,399938	0,599994	0,499878
Истинное значение	0,4	0,6	0,5

рошо. Успех метода решающим образом зависит от начального приближения. Другими словами, мы можем надеяться на успешное применение метода, только если вначале мы имеем по крайней мере грубое представление о природе действительного физического процесса. (Программы вычислений приведены в приложении 3.)

В случае (в) с $N=7$ метод квазилинеаризации требует интегрирования 124 линейных дифференциальных уравнений и решений системы 3 линейных алгебраических уравнений. Расчеты занимают около двух минут на машине IBM-7044.

Впоследствии мы укажем, как комбинация динамического программирования и квазилинеаризации приводит к лучшему методу определения положения внутренних поверхностей разделов в слоистой среде.

8. Уравнение Ван-дер-Поля

Одним из наиболее известных нелинейных дифференциальных уравнений является уравнение Ван-дер-Поля

$$x'' + \lambda(x^2 - 1)x' + x = 0. \tag{8.1}$$

Хотя эквивалентное ему нелинейное уравнение было исследовано уже Рэлеем, данное уравнение получило известность только в связи с развитием современной электроники. Это уравнение превосходно описывает поведение мультивибратора и поэтому находит приложение во многих других областях науки, например в кардиологии и электроэнцефалографии, в которых наблюдаются явления «переброса». На это указал сам Ван-дер-Поль в своей статье.

Определенный интерес представляет вопрос, можно ли по результатам наблюдения за решением эффективно определить

параметр λ. Пусть сделаны следующие три наблюдения смещения x:

$$\begin{aligned} x(4) &= -1{,}80843, \\ x(6) &= -1{,}63385, \\ x(8) &= -1{,}40456. \end{aligned} \tag{8.2}$$

Мы хотим определить λ и $u(4)=x'(4)$. В использовании метода наименьших квадратов нет необходимости, поскольку задано только три наблюдения. Для получения начального приближения заметим, что

$$\begin{aligned} \frac{x(6)-x(4)}{2} &\cong 0{,}087 \cong x'(4), \\ \frac{x(8)-x(6)}{2} &\cong 0{,}114 \cong x'(6), \\ \frac{x'(6)-x'(4)}{2} &\cong 0{,}014 \cong x''(4), \end{aligned} \tag{8.3}$$

Отсюда с помощью (8.1) получаем

$$\lambda_0 = 7 \tag{8.4}$$

как начальное приближение для λ. Затем мы интегрируем уравнение (8.1) или эквивалентную ему систему

$$\begin{aligned} x' &= u, \\ u' &= -\lambda(x^2-1)u - x \\ \lambda' &= 0 \end{aligned} \tag{8.5}$$

с начальными условиями

$$\begin{aligned} x(4) &= -1{,}80843, \\ u(4) &= +0{,}08, \\ \lambda(4) &= 7{,}0. \end{aligned} \tag{8.6}$$

Чтобы получить $(n+1)$-е приближение после вычисления n-го, используем линеаризованные соотношения

$$\begin{aligned} \dot{x}_{n+1} &= u_{n+1}, \\ \dot{u}_{n+1} &= -\lambda_n(x_n^2-1)u_n - x_n + (x_{n+1}-x_n)[-2\lambda_n x_n u_n - 1] + \\ &\quad + (u_{n+1}-u_n)[-\lambda_n(x_n^2-1)] + (\lambda_{n+1}-\lambda_n)[-(x_n^2-1)u_n], \\ \dot{\lambda}_{n+1} &= 0 \end{aligned} \tag{8.7}$$

вместе с граничными условиями в трех точках (8.2).

Таблица 6.4

Первый численный эксперимент

	Начальное приближение	Итерация 2	Итерация 3	Итерация 4	Истинные значения
x (4)	—1,80843	—1,80843	—1,80843	—1,80843	—1,8084322
u (4)	+0,08	+0,0564454	+0,0794911	+0,079366	0,079366909
λ (4)	+7,0	9,91541	10,0004	10,00000	10,00000

Результаты численного эксперимента приведены в табл. 6.4. Второй эксперимент проводился с худшими начальными приближениями скорости и параметра системы. Эти результаты

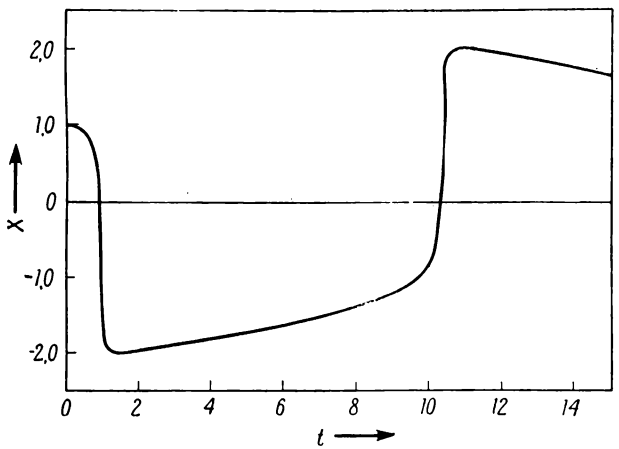


Рис. 6.3. Решение уравнения Ван-дер-Поля при $\lambda = 10{,}0$, $x(0) = 1{,}0$, $\dot{x}(0) = 0{,}0$.

представлены в табл. 6.5. Третий эксперимент, в котором начальное приближение для параметра системы бралось равным 20, привел к переполнению и, таким образом, не дал результатов.

Таблица 6.5

Второй численный эксперимент

	Начальное приближение	Итерация 2	Итерация 4	Итерация 6	Истинные значения
x (4)	—1,80843	—1,80843	—1,80843	—1,80843	—1,8084322
u (4)	0,1000	—2,0758	0,599288	0,0792063	0,079366909
λ (4)	5,000	3,87992	11,4091	9,99956	10,000

Значения (8.2) были получены путем интегрирования уравнения Ван-дер-Поля при

$$\lambda = 10{,}0, \quad x(0) = 1{,}0, \quad \dot{x}(0) = 0{,}0. \tag{8.8}$$

График представлен на рис. 6.3. (Программа вычислений приведена в приложении 4.)

9. Определение орбиты как многоточечная краевая задача

Общая задача решения векторного уравнения

$$x' = f(x, t), \quad 0 \leqslant t \leqslant b, \tag{9.1}$$

подчиненного многоточечным краевым условиям

$$\sum_{j=1}^{N} a_{ij} x_j(t_i) = b_i, \quad i = 1, 2, \ldots, N, \tag{9.2}$$

где $0 \leqslant t_1 \leqslant t_2 \leqslant \ldots \leqslant t_N \leqslant b$, уже упоминалась ранее в связи с квазилинеаризацией. Сейчас мы рассмотрим довольно интересный случай, встречающийся в задаче определения орбиты.

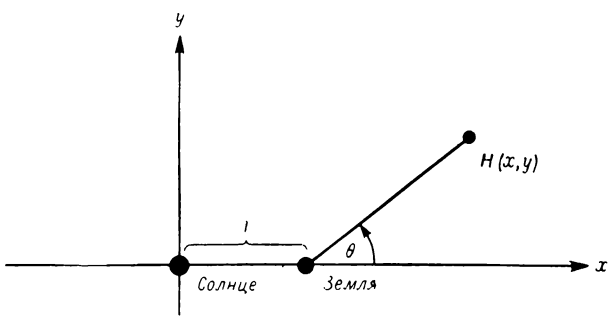


Рис. 6.4. Физическая схема.

Пусть уравнения движения небесного тела в безразмерных переменных заданы в виде

$$x' = u, \quad y' = v, \quad u' = -\frac{x}{(x^2+y^2)^{1,5}}, \quad v' = -\frac{y}{(x^2+y^2)^{1,5}}. \tag{9.3}$$

Движение происходит в плоскости x, y — плоскости эклиптики. В моменты t_i, $i = 1, 2, 3, 4$, заданы величины $\theta(t_i) = \theta_i$, по-

рождающие условия

$$y(t_i) = (x(t_i) - 1)\operatorname{tg}\theta_i, \quad i = 1, 2, 3, 4 \tag{9.4}$$

(рис. 6.4).

Мы хотим определить орбиту, т. е. получить в некоторый момент времени, например при $t=2{,}5$, значения x, y, u, v. Для простоты предположим, что Земля закреплена в точке (1,0) и что единственной силой, действующей на наше небесное тело, является гравитационное притяжение Солнца [1]).

10. Численные результаты

Рассмотрим случай, для которого данные наблюдений (в радианах) равны

$$\begin{aligned} &\theta(0{,}5) = 0{,}251297, \quad \theta(1{,}0) = 0{,}510240, \\ &\theta(1{,}5) = 0{,}783690, \quad \theta(2{,}0) = 1{,}076540. \end{aligned} \tag{10.1}$$

Эти данные получены в предположении, что в нулевой момент времени $x(0)=2{,}0$, $y(0)=0$, $u(0)=0$, $v(0)=0{,}5$. Общая схема (с неподвижной Землей) изображена на рис. 6.5. Траекторию начального приближения выберем следующим образом: пусть в нулевой момент времени наблюдаемое тело находится в точке (1,0), и пусть оно движется по круговой орбите с периодом 3,664. Это очень плохое начальное приближение. Затем, пользуясь методом интегрирования Рунге — Кутта и процедурой квазилинеаризации, мы получаем результаты, приведенные в табл. 6.6.

Таблица 6.6

Приближения координат и скоростей в момент времени $t = 2{,}5$

	Начальное приближение	Первое приближение	Второе приближение	Третье приближение	Четвертое приближение	Точное значение до шести значащих цифр
$x(2{,}5)$	—0,455175	1,554360	1,169660	1,190170	1,193560	1,193610
$u(2{,}5)$	—0,873320	—1,638830	—0,495101	—0,656686	—0,664207	—0,664263
$y(2{,}5)$	0,975790	2,982690	0,865184	1,043930	1,060540	1,060700
$v(2{,}5)$	—0,408690	1,807620	0,201206	0,237695	0,247433	0,247499

[1]) Это предположение соответствует переходу к системе координат, вращающейся вместе с Землей. Но тогда в правые части уравнений (9.3) для u и v следует добавить дополнительные члены — центробежные и кориолисовы силы инерции, обусловленные вращением системы координат. — *Прим. ред.*

Расчеты на IBM-7090 заняли примерно полторы минуты, причем не предпринималась попытка какой-либо рационализации вычислений. Разумеется, получены не только результаты табл. 6.6, но и промежуточные данные для $0 \leqslant t \leqslant 2{,}5$ с интервалом 0,01. (Программы приведены в приложении 5.)

Удивительно, что если в качестве начального приближения к орбите выбрать точку, совпадающую с положением Земли, мы

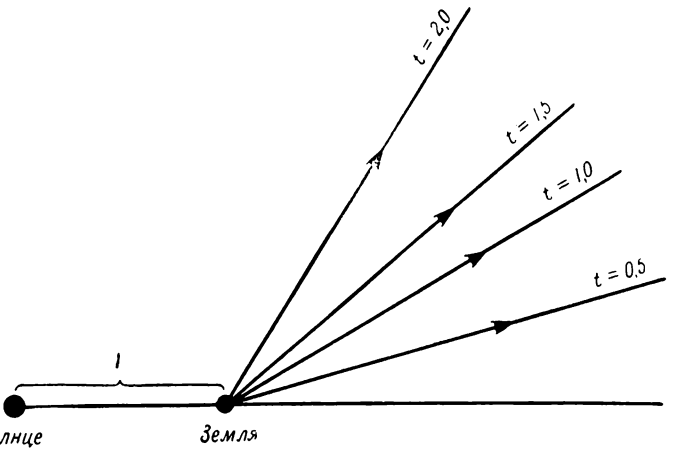


Рис. 6.5. Приближенные направления на небесное тело в четыре момента времени наблюдений с неподвижной Земли.

все же получим такую же быструю сходимость к истинной орбите.

Этот метод легко обобщить с помощью метода наименьших квадратов на случай, когда известно более четырех наблюдений. Кроме того, возмущения, возникающие вследствие движения Земли и наличия других планет, легко учитываются, поскольку они приводят просто к изменениям уравнений (9.3).

11. Анализ периодических зависимостей

Одной из основных научных проблем является проблема выявления скрытых периодичностей. В самом деле, трудно представить себе область, в которой самым серьезным образом не возникали бы вопросы такой природы. В аналитическом виде проблема формулируется очень просто. Пусть, зная о том, что функция $f(t)$ имеет вид

$$f(t) = \sum_{i=1}^{R} a_i \cos \omega_i t, \tag{11.1}$$

и располагая рядом наблюдений $b_i = f(t_i)$, $i = 1, 2, \ldots, M$, мы хотим определить амплитуды и частоты, а также число R. Рассмотрим более простую задачу, когда R считается известным, и возьмем $M > 2R$.

Наш подход, отличающийся от любого стандартного подхода, состоит в том, чтобы превратить эту задачу в вариационную задачу с многоточечными граничными условиями и затем использовать квазилинеаризацию. Мы исходим из выражения

$$f(t) = \sum_{i=1}^{R} u_i(t), \tag{11.2}$$

где u_i — решения уравнений

$$u_i'' + \omega_i^2 u_i = 0, \quad u_i(0) = \alpha_i, \quad u_i'(0) = 0. \tag{11.3}$$

Наша цель — определить α_i и ω_i, минимизирующие величину

$$\sum_{i=1}^{M} (f(t_i) - b_i)^2. \tag{11.4}$$

При исследовании мы используем, как и выше, квазилинеаризацию, исходя из начального приближения $(\alpha_i^{(0)}, \omega_i^{(0)})$, порождающего начальные значения функций $u_i^{(0)}(t)$. В следующем разделе мы приведем некоторые численные результаты.

12. Численные результаты

Рассмотрим простой случай:

$$f(t) = \sum_{i=1}^{3} \alpha_i \cos \omega_i t, \tag{12.1}$$

где

$$\begin{aligned} \alpha_1 &= 1{,}000000, \quad \omega_1 = 1{,}11000, \\ \alpha_2 &= 0{,}500000, \quad \omega_2 = 2{,}03000, \\ \alpha_3 &= 0{,}100000, \quad \omega_3 = 3{,}42000. \end{aligned} \tag{12.2}$$

Пусть проведено шесть наблюдений:

t	$f(t)$
0,00	1,60000
1,00	0,126895
2,00	—0,823201
3,00	—0,558708
4,00	—0,356338
5,00	0,351082

Результаты последовательных приближений собраны в табл. 6.7.

Таблица 6.7

Приближение	i	α_i	ω_i
0	1	0,900000	1,00000
	2	0,600000	2,00000
	3	0,200000	3,00000
1	1	0,943582	1,09000
	2	0,486676	1,92583
	3	0,169742	2,45651
2	1	1,00218	1,11113
	2	0,476529	2,01352
	3	0,121287	2,62338
3	1	1,00002	1,11001
	2	0,503119	2,03179
	3	0,096864	2,80801
4	1	0,999988	1,11000
	2	0,499537	2,02981
	3	0,100475	2,85931
Истинное значение	1	1,000000	1,11000
	2	0,500000	2,03000
	3	0,100000	3,42000

Полученные результаты для случая семи наблюдений

t	$f(t)$
0,00	1,60000
0,83	0,452413
1,67	—0,679691
2,50	—0,820313
3,33	—0,367504
4,17	—0,381874
5,00	0,351082

Таблица 6.8

Приближение	i	α_i	ω_i
0	1	0,900000	1,00000
	2	0,600000	2,00000
	3	0,200000	3,00000
1	1	0,984727	1,14378
	2	0,535234	2,07572
	3	0,057545	3,29235
2	1	0,998859	1,11124
	2	0,491950	2,02810
	3	0,107930	3,39973
3	1	1,00904	1,11003
	2	0,500095	2,03018
	3	0,099857	3,41930
4	1	1,000000	1,11000
	2	0,500000	2,03000
	3	0,100000	3,42000
Истинное значение	1	1,000000	1,11000
	2	0,500000	2,03000
	3	0,100000	3,42000

приведены в табл. 6.8, а для восьми наблюдений

t	$f(t)$
0,00	1,60000
0,71	0,694147
1,43	—0,484599
2,14	—0,849519
2,86	—0,649072
3,57	—0,302561
4,29	—0,378655
5,00	0,351082

в табл. 6.9.

Таблица 6.9

Приближение	i	α_i	ω_i
0	1	0,900000	1,00000
	2	0,600000	2,00000
	3	0,200000	3,00000
1	1	0,981747	1,13890
	2	0,531127	2,06075
	3	0,056783	3,22805
2	1	0,998520	1,11046
	2	0,492646	2,02508
	3	0,105177	3,41224
3	1	1,00006	1,11003
	2	0,500003	2,03016
	3	0,099888	3,41981
4	1	1,000000	1,11000
	2	0,500000	2,03000
	3	0,100000	3,42000
Истинное значение	1	1,000000	1,11000
	2	0,500000	2,03000
	3	0,100000	3,42000

13. Обсуждение

Выяснилось, что для получения сходимости требуется довольно хорошее начальное приближение. Например, когда экспериментальные расчеты, подобные проиллюстрированным в табл. 6.7 и 6.8, были проведены с начальными значениями

$$\alpha_1 = \alpha_2 = \alpha_3 = 1{,}0, \qquad \omega_1 = 7, \quad \omega_2 = 6, \quad \omega_3 = 8, \tag{13.1}$$

оказалось, что последовательные приближения не сходятся.

14. Периодические вынуждающие члены

Аналогичные методы можно использовать для определения α_i и ω_i в вынуждающих членах, когда наблюдаемая функция $u(t)$ удовлетворяет линейному дифференциальному уравнению вида

$$u'' + b_1 u' + b_2 u = \alpha_1 \cos \omega_1 t + \alpha_2 \cos \omega_2 t, \qquad u(0) = c_1, \quad u'(0) = c_2. \tag{14.1}$$

На самом деле можно действовать намного лучше. Имея ряд наблюдений $\{u(t_i)\}$, можно определить начальные условия c_1, c_2, параметры системы b_1, b_2 и параметры вынуждающей функции в правой части уравнения (14.1).

15. Оценка параметров сердца

В последнее время появилась возможность дать объяснение потенциалам, наблюдаемым на коже пациента и возникающим благодаря деполяризации желудочков сердца. Основная идея

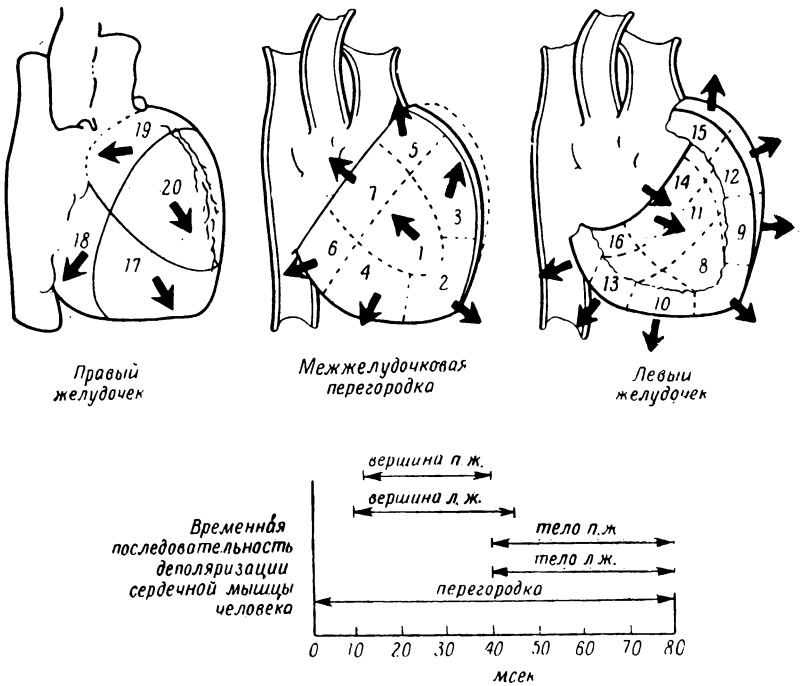
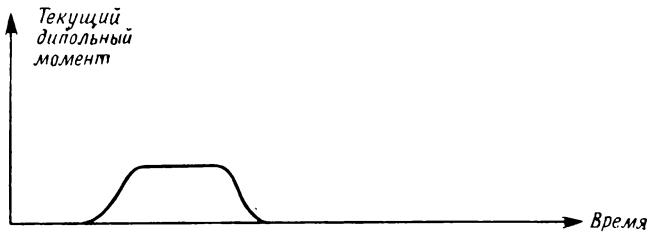


Рис. 6.6. Сегменты сердечной мышцы.

состоит в разделении желудочков на ряд сегментов, каждый из которых содержит дипольный источник тока. Диполь расположен в центре сегмента, и его ось ортогональна поверхности сегмента. Диполи обладают изменяющимися во времени моментами, зависящими от деполяризационной волны, бегущей по желудочкам. Эта модель проиллюстрирована на рис. 6.6 и 6.7. Постулированная последовательность деполяризационных фаз получена путем экстраполяции хорошо известных экспериментальных результатов А. Шера для сердца собаки. Тело можно

рассматривать как объемный проводник. Таким образом, можно получить изменение во времени поверхностного потенциала, которое, как показано многочисленными расчетами на быстродействующей счетной машине, хорошо согласуется с результатами, полученными для нормального сердца и сердца с отклонениями от нормы.

Цель настоящего обсуждения — показать, что возможно также решение обратной задачи: задавая изменение во времени



Р и с. 6.7. Типичная зависимость дипольного момента от времени $F(t)$.

поверхностных потенциалов, можно получить изменение во времени отдельных дипольных моментов. С математической точки зрения мы рассматриваем эту задачу как нелинейную многоточечную краевую задачу. Она решается с помощью квазилинеаризации в соответствии с идеями, изложенными на предыдущих страницах.

16. Основные предположения

Введем стандартную ортогональную систему координат x, y, z и рассмотрим диполь i-го сегмента, расположенный в точке (x_i, y_i, z_i), причем l_i, m_i, n_i — направляющие косинусы его оси. Этот диполь создает в точке (a_j, b_j, c_j) потенциал w, зависящий от свойств рассматриваемого тела и окружающей среды как объемных проводников. Если для простоты рассмотреть однородную среду бесконечной протяженности, то можно записать

$$w(a_j, b_j, c_j, t) = (kF_i(t)\cos\theta_{ij})\, r_{ij}^{-2}, \tag{16.1}$$

где

$$\begin{gathered} k = \text{const}, \\ F_i(t) — \text{дипольный момент } i\text{-го сегмента} \\ \text{в момент времени } t, \\ \cos\theta_{ij} = [l_i(a_j - x_i) + m_i(b_j - y_i) + n_i(c_j - z_i)]\, r_{ij}^{-1}, \\ r_{ij}^2 = (a_j - x_i)^2 + (b_j - y_i)^2 + (c_j - z_i)^2. \end{gathered} \tag{16.2}$$

Точки, бесконечно удаленные от диполей, имеют нулевой потенциал. Введем опорную точку (a_0, b_0, c_0) и предположим, что относительно нее в N точках на поверхности определены разности потенциалов. Кроме того, предположим, что желудочки разделены на M сегментов.

Теперь мы должны подробнее рассмотреть величины дипольных моментов $F_i(t)$, $i=1, 2, \ldots, M$, как функции времени. Одна из таких функций показана на рис. 6.7. Мы выяснили, что такие кривые хорошо представляются решением обыкновенных дифференциальных уравнений

$$\frac{dF_i}{dt} = k_i(t_i - t)F_i, \quad F_i(0) = h_i, \quad 0 \leqslant t \leqslant 80 \text{ мсек}, \quad (16.3)$$
$$i = 1, 2, \ldots, M.$$

Параметры t_i — времена, в которые достигаются максимумы моментов, k_i связаны с шириной кривых, а h_i с максимумами моментов[1]). Аналитическое решение этих уравнений нас не интересует, так как при более детальных исследованиях уравнения (16.3) будут заменяться на более сложные дифференциальные уравнения.

Обозначим через $V_j(t)$, $j=1, 2, \ldots, N$, потенциал, возникающий в момент времени t $(0 \leqslant t \leqslant 80$ *мсек*$)$ в j-й наблюдаемой точке. Тогда

$$V_j(t) = k \sum_{i=1}^{M} (F_i(t) \cos \theta_{ij}) r_{ij}^{-2} - \sum_{i=1}^{M} (F_i(t) \cos \theta_i) r_i^{-2}, \quad j = 1, 2, \ldots, N, \quad (16.4)$$

где

$$\cos \theta_i = [l_i(a_0 - x_i) + m_i(b_0 - y_i) + n_i(c_0 - z_i)] r_i^{-1},$$
$$r_i^2 = (a_0 - x_i)^2 + (b_0 - y_i)^2 + (c_0 - z_i)^2, \quad (16.5)$$
$$i = 1, 2, \ldots, M.$$

17. Обратная задача

Предположим, что сердце наблюдалось в течение времени $0 \leqslant t \leqslant 80$ *мсек* и что записана временная последовательность $\{b_j(t)\}$ поверхностных потенциалов, а именно:

$b_j(t)$ — наблюдаемый поверхностный потенциал в момент времени t и в j-й точке наблюдения, $j = 1, 2, \ldots, N$ и $t = 0, 1, 2, \ldots, 80$ *мсек*.

[1]) Это соответствует аппроксимации вида $F_i(t) = C_i \exp[-k_i(t - t_i)^2/2]$, где C_i, t_i, k_i — постоянные. — *Прим. ред.*

Мы хотим определить дипольные моменты $F_i(t)$, $i=1, 2, \ldots, M$, которые порождают поверхностные потенциалы, наилучшим образом совпадающие с наблюдаемыми. Переведем это на более точный математический язык. Мы хотим определить $3M$ констант k_i, t_i и h_i в дифференциальных уравнениях (16.3) таким образом, чтобы достигался минимум суммы квадратов разностей между вычисленными и наблюдаемыми потенциалами в точках наблюдения. Путем соответствующего выбора параметров сердца t_i, k_i и h_i, $i=1, 2, \ldots, M$, мы хотим минимизировать функцию

$$S=\sum_{r=0}^{80}\sum_{j=1}^{N}(V_j(r)-b_j(r))^2, \tag{17.1}$$

где $V_j(t)$ определяются из (16.4), а $F_i(t)$ определяются из (16.3).

18. Численные эксперименты

Предварительная вычислительная работа убедила нас в том, что при решении обратной задачи можно действовать так, как описано в предыдущем разделе. Для того чтобы подтвердить это, было проведено несколько численных экспериментов. Сначала мы выбрали число M сегментов желудочков, равное пяти, а число N наблюдаемых точек, равное трем. Затем параметрам t_i, k_i и h_i, $i=1, 2, \ldots, 5$, придавались различные значения с тем, чтобы промоделировать нормальное сердце, а также аномальные сердца (гипертрофии и инфаркты). Кроме того, задавались значения (x_i, y_i, z_i), (a_i, b_i, c_i) и (l_i, m_i, n_i). Интегрируя уравнения (16.3) с этими значениями и используя (16.4) и (16.5), мы получили наборы гипотетических поверхностных потенциалов, отвечающих различным сердцам. Мы использовали сегменты 1, 8, 11, 14 и 20, и эти пять сегментов сердца дали хорошее представление о потенциалах, измеренных тремя прямоугольными пластинами у нормальных людей (рис. 6.8).

Мы предполагали, что известны значения t_i, $i=1, 2, \ldots, 5$, которые мы выбираем для представления сердца как обычной проводящей системы. Мы предполагали также, что имеется набор измерений поверхностных потенциалов $b_j(r)$, $j=1, 2, 3$, $r=0, 1, 2, \ldots, 80$. Затем с помощью метода квазилинеаризации были получены «недостающие» значения параметров сердца k_i и h_i, $i=1, 2, \ldots, 5$. Следует иметь в виду, что объем вычислительной работы значителен. Требуется решить систему из 110 линейных дифференциальных уравнений, причем на каждом этапе вычислений требуется решать систему десяти линейных алгебраических уравнений. Многое еще нужно сделать для ис-

следования области сходимости метода, для определения влияния ошибок наблюдения поверхностных потенциалов, а также для максимально возможного повышения эффективности вычислительных программ. Изложенные предварительные эксперименты просто указывают на возможность данного подхода

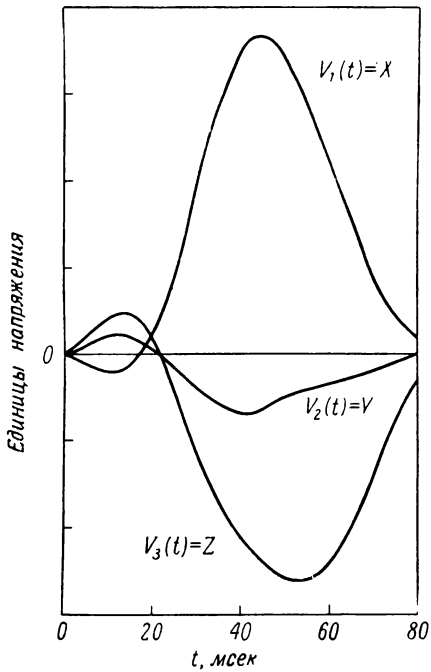


Рис. 6.8. Потенциалы, полученные из пятисегментной модели нормального сердца.

к постановке и численному решению обратных задач кардиологии. (Несколько программ дано в приложении 6.)

КОММЕНТАРИИ И БИБЛИОГРАФИЯ

2. См.

Bellman R., Kagiwada H., Kalaba R., A computational procedure for optimal system design and utilization, *Proc. Nat. Acad. Sci. USA*, **48** (1962), 1524—1528;

* Понтрягин Л. С., Болтянский В. Г., Гамкрелидзе Р. В., Мищенко Е. Ф., Математическая теория оптимальных процессов, Физматгиз, М., 1961.

5. Этот материал был впервые изложен в работах

Bellman R., Kagiwada H., Kalaba R., Ueno S., Inverse problems in radiative transfer: layered media, The RAND Corporation, RM-4281-ARPA, December 1964 (в печати в *Icarus*);
Bellman R., Kagivada H., Kalaba R., Ueno S., On the identification of systems and the unscrambling of data — II: An inverse problem in radiative transfer, The RAND Corporation, RM-4332-ARPA, November 1964.

См. также

* Bellman R., Kalaba R., Prestrud M., Imbedding and radiative transfer in slabs of finite thickness, Modern Analytic and Computational Methods in Science and Mathematics, vol. I, New York, 1963;
* Bellman R., Kagiwada H., Kalaba R., Prestrud M., Invariant imbedding and time-dependent transport processes, Modern Analytic and Computational Methods in Science and Mathematics, vol. II, New York, 1964.

Ниже приводится таблица элементов матрицы отражения при $N=7$ для слоя толщины $\tau_1=1{,}0$, функция альбедо $\lambda(\tau)$ равна $0{,}5+0{,}1\text{th}[10(\tau-0{,}5)]$.

Измерения $\{b_{ij}\}$

	$j=1$	2	3	4	5	6	7
$j=1$	0,079914	0,28164	0,014304	0,009104	0,006707	0,005515	0,004970
2	0,143038	0,091522	0,058437	0,040826	0,031405	0,026378	0,023989
3	0,167000	0,134331	0,099653	0,075106	0,060044	0,051445	0,047248
4	0,178898	0,157955	0,126408	0,099392	0,081253	0,070435	0,065042
5	0,185284	0,170817	0,142072	0,114229	0,094495	0,082423	0,076332
6	0,188723	0,177733	0,150791	0,122665	0,102104	0,089349	0,082870
7	0,190354	0,180898	0,154995	0,126773	0,105829	0,092748	0,086083

Интересно сравнить эти данные с соответствующими табличными значениями, приведенными в томе I из упомянутых выше книг для случая однородных слоев толщины 1,0 и величин альбедо 0,4; 0,5 и 0,6. Имеются значительные различия вследствие неоднородности.

8. Наше обсуждение опирается на работу

Bellman R., Kagiwada H., Kalaba R., Quasilinearization, system identification and prediction, The RAND Corporation, RM-3812-PR, August 1963.

Эти методы находят интересное применение в электроэнцефалографии. О постановке задачи см.

Dewan E. M., Nonlinear oscillations and electroencephalography, *J. Theoret. Biol.*, 7 (1964), 141—159.

9. Этот подход намечен в работе

Bellman R., Kagiwada H., Kalaba R., Orbit determination as a multipoint boundary value problem and quasilinearisation, *Proc. Nat. Acad. Sci. USA*, 48 (1962), 1327—1329.

Интересное продвижение в области программирования при определении частных производных сделано в работе

Bellman R., Kagiwada H., Kalaba R., Wengert's numerical method for partial derivatives, orbit determination and quasilinearization, The RAND Corporation, RM-4354-PR, November 1964 (в печати в *Comm. ACM*).

О другом приложении см.

Lavi A., Strauss J. C., Parameter identification in continuous systems (в печати).

11. В основе этого раздела лежит работа

Bellman R., Kagiwada H., Kalaba R., On the identification of systems and the unscrambling of data — I: Hidden periodicities, The RAND Corporation, RM-4285-PR, September 1964.

15. Математическая модель представлена в работе

Selvester R., Collier C., Peterson R., Analogue computer model of the vector-cardiogram, *Circulation Research* (в печати).

Этот раздел основан на исследованиях, изложенных в статье

Bellman R., Collier C., Kagiwada H., Kalaba R., Selvester R., Estimation of heart parameters using skin potential measurements, *Comm. ACM*, 7 (1964), 666—668.

Мы уклонились от какого-либо обсуждения современных приложений комбинации динамического программирования и квазилинеаризации к задачам идентификации систем. По этому вопросу см.

Bellman R., Gluss B., Roth R., Identification of differential systems with time-varying coefficients, The RAND Corporation, RM-4288-PR, November 1964;
Bellman R., Gluss B., Roth R., Segmental differential approximation and the «Black box» problem, *J. Math. Anal. Appl.* (в печати);
Bellman R., Gluss B., Roth R., On the identification of systems and the unscrambling of data: Some problems suggested by neurophysiology, *Proc. Nat. Acad. Sci. USA*, **52** (1964), 1239—1240.

О приложении преобразования Лапласа см.

Bellman R., Kagiwada H., Kalaba R., Identification of linear systems using numerical inversion of Laplace transforms, *IEEE Trans.* (в печати).

Глава VII

Динамическое программирование и квазилинеаризация

1. Введение

В этой главе мы хотим обсудить некоторые связи динамического программирования с теорией положительных операторов и квазилинеаризацией. Сначала проследим истоки квазилинеаризации в концепции приближения в пространстве политик. Вслед за этим укажем, как можно использовать комбинацию квазилинеаризации и динамического программирования для получения приближенного метода решения многомерных вариационных задач, — это значительно расширяет область применения обоих подходов. Кроме того, мы покажем, как сочетать динамическое программирование и квазилинеаризацию для рассмотрения некоторых сложных задач идентификации.

Как и в предыдущих главах, ссылки на детальные изложения читатель найдет в конце главы.

2. Основное функциональное уравнение

Рассмотрим многошаговый процесс решения в пространстве состояний S, где $p \in S$ представляет некое типичное состояние. Пусть $T(p, q)$ — элемент семейства преобразований пространства S, характеризуемый вектором политики $q \in D$, где D есть пространство политик или стратегий, и пусть $g(p, q)$ — доход на одном шаге.

Если через $f(p)$ обозначить полный доход в бесконечношаговом процессе, полученный использованием оптимальной политики, то принцип оптимальности приведет к функциональному уравнению

$$f(p) = \max_q [g(p, q) + f(T(p, q))]. \tag{2.1}$$

Сразу же возникает интересный вопрос об установлении существования и единственности решения этого уравнения при различных предположениях относительно функций g и T и пространств S и D.

Для этой цели можно привлечь ряд подходов классического анализа, таких, как теоремы о неподвижной точке, аппроксимация дискретными процессами и метод последовательных приближений. Так, например, используя последний из названных

методов, можно изучить уравнение (2.1) при помощи последовательности $\{f_N(p)\}$, порожденной соотношением

$$f_N(p) = \max_q [g(p, q) + f_{N-1}(T(p, q))], \quad N \geqslant 1, \tag{2.2}$$

где $f_0(p)$ выбрано некоторым специальным образом.

Этот метод приближения имеет большое значение, поскольку он эквивалентен рассмотрению неограниченного процесса как предела конечно-шаговых процессов.

3. Приближение в пространстве политик

Одно из преимуществ теории динамического программирования состоит в том, что она допускает и фактически рекомендует тип приближений, называемый приближением в пространстве политик, который отсутствует в классическом анализе. Основная идея его очень проста. Вместо аппроксимации функции $f(p)$ мы аппроксимируем функцию политики $q(p)$. С точки зрения приложений это очень естественный подход. Как мы сейчас увидим, он обладает важным аналитическим достоинством.

Пусть задана начальная политика $q_0(p)$. Определим соответствующую функцию дохода $f_0(p)$ с помощью функционального уравнения

$$f_0(p) = g(p, q_0) + f_0(T(p, q_0)). \tag{3.1}$$

Чтобы улучшить $q_0(p)$, найдем функцию $q_1(p)$, максимизирующую выражение

$$g(p, q) + f_0(T(p, q)). \tag{3.2}$$

Используя политику $q_1(p)$, получаем функцию $f_1(p)$ как решение уравнения

$$f_1(p) = g(p, q_1) + f_1(T(p, q_1)). \tag{3.3}$$

Как и для уравнения (3.1), решение получается прямой итерацией:

$$f_1(p) = g(p, q_1) + g(T(p, q_1), q_1(T(p, q_1))) + \ldots. \tag{3.4}$$

Таким образом, мы получаем последовательность политик $\{q_N(p)\}$ и функций дохода $\{f_N(p)\}$.

Теперь изучим вопрос, при каких условиях обосновано утверждение о том, что $q_1(p)$ есть улучшение $q_0(p)$. Мы имеем

$$f_0(p) = g(p, q_0) + f_0(T(p, q_0)) \leqslant g(p, q_1) + f_0(T(p, q_1)) \tag{3.5}$$

в силу способа, которым выбирается q_1. Сравнивая это неравенство с уравнением (3.3) для $f_1(p)$, мы видим, что

неравенство $f_0 \leqslant f_1$ является следствием свойства положительности оператора [1])

$$T(f) = f(p) - f(T(p, q)). \tag{3.6}$$

Таким образом, приближение в пространстве политик приводит нас в область функциональных неравенств. Убедившись однажды в преимуществах этого подхода, нетрудно сделать шаг к методике, изложенной в предыдущих главах.

Предполагая, что $T(f)$ обладает требуемыми свойствами положительности, покажем, что каждая функция $f_n(p)$ ограничена функцией $f(p)$. Имеем

$$\begin{aligned} f_n(p) &= g(p, q_n) + f_n(T(p, q_n)), \\ f(p) &= \max_q [g(p, q) + f(T(p, q))] \geqslant g(p, q_n) + f(T(p, q_n)), \end{aligned} \tag{3.7}$$

откуда следует искомое неравенство.

Монотонность сходимости дает очень гибкий метод доказательства существования решения уравнения (2.1).

4. Динамическое программирование и квазилинеаризация

Основная трудность применения теории квазилинеаризации состоит в том, что для обеспечения сходимости последовательных приближений требуется выполнение жестких условий. Однако приведенные в гл. VI примеры показывают, что эти методы часто на практике работают лучше, чем можно было бы ожидать. Тем не менее состояние дел не очень удовлетворительное, и весьма важна разработка простых и надежных (гарантирующих сходимость) методов.

Предположим, например, что мы хотим использовать квазилинеаризацию в задаче минимизации функционала

$$J(u) = \int_0^T g(u, u')\,dt, \tag{4.1}$$

где u — скалярная функция, подчиненная краевым условиям $u(0) = c_0$, $u(T) = c'$. Если в качестве начального приближения $u_0(t)$ взять прямую, соединяющую точки $(0, c_0)$ и (T, c'), то для того, чтобы гарантировать сходимость, мы должны ограничить T. Если начальное приближение выбрано более искусно, то можно допустить для T бóльшие значения. К сожалению, во

---

[1]) Фактически здесь используется следующее свойство оператора T: неравенство $T(f_0) \leqslant T(f_1)$ влечет за собой неравенство $f_0 \leqslant f_1$. — *Прим. ред.*

многих приложениях мы не обладаем достаточной информацией для того, чтобы выбрать начальное приближение, лучшее, чем простое приближение того типа, о котором говорилось выше.

Для преодоления этой трудности используем сочетание квазилинеаризации и динамического программирования. По существу, мы собираемся применить квазилинеаризацию для *локальной* оптимизации, а динамическое программирование для *глобальной* оптимизации.

Чтобы проиллюстрировать комбинацию методов, вернемся к поставленной выше вариационной задаче. На плоскости c, T мы строим сетку из точек, как показано на рис. 7.1.

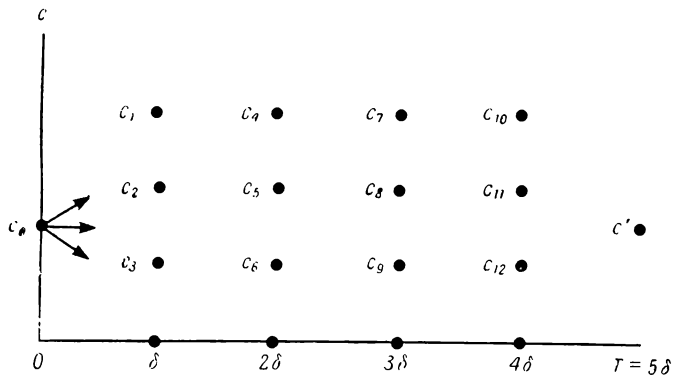


Рис. 7.1.

Мы взяли небольшое число точек (и одномерный пример) для того, чтобы пояснить идею, не усложняя чертежа или обозначений. На практике при рассмотрении многомерных вариационных задач можно легко построить в фазовом пространстве сетки из нескольких тысяч клеток.

Предположим, что из точки c_0 можно двигаться только в одну из трех точек c_1, c_2, c_3; из c_1, c_2, c_3 — в одну из точек c_4, c_5, c_6 и т. д.; из c_{10}, c_{11} или c_{12} нужно идти в c'. Допустимыми траекториями из одной точки сетки в другую являются те, которые минимизируют функционал

$$J(u) = \int_{k\delta}^{(k+1)\delta} g(u, u')dt, \tag{4.2}$$

где $u(k\delta) = c_i$, $u[(k+1)\delta] = c_j$, а c_i и c_j — «ближайшие соседи» в определенном выше смысле. Назовем эти траектории *геодезическими*, а соответствующее значение $J(u)$ *геодезической длиной*. Обозначим эту длину через d_{ij}.

Выше мы указывали, как можно определить с помощью квазилинеаризации геодезические расстояния, если δ достаточно мало. Границу величины δ можно найти априори через оценки для функции g и ее частных производных. Сейчас мы покажем, как с помощью динамического программирования соединить эти геодезические дуги и тем самым найти превосходное приближение к действительной минимизирующей функции.

5. Функциональные уравнения

Поставленная в предыдущем разделе задача оптимизации является частным случаем более общей транспортной задачи, для которой существуют очень эффективные алгоритмы. В данном случае структура сети делает вычисление особенно простым.

Введем функцию

$$f_i — \text{минимум полной геодезической длины от } c_i \text{ до } c', \tag{5.1}$$

определенную для $i=0, 1, \ldots, 12$. Тогда принцип оптимальности дает функциональное уравнение

$$f_i = \min_{j \neq i} [d_{ij} + f_j], \tag{5.2}$$

связывающее величины f_i. В данном случае это уравнение сводится к рекуррентному соотношению, так как f_i немедленно определяются для точек c_{10}, c_{11}, c_{12}. Для них мы имеем

$$f_i = d_{i,\,c'}. \tag{5.3}$$

Используя эти значения в уравнении (5.2), можно вычислить f_i для c_7, c_8, c_9. Продолжая дальше действовать так же, мы производим вычисления для точек c_4, c_5, c_6 и c_1, c_2, c_3 и, наконец, в точке c_0 получаем искомое значение функции.

6. Обсуждение

Как только величины d_{ij} определены, вычисление f_i требует незначительного времени. Время, необходимое для определения величин d_{ij}, разумеется, зависит непосредственно от их числа, которое в свою очередь зависит от δ. Однако имеется много приемов, которые можно использовать для уменьшения времени вычислений.

Определив таким способом оптимальную «ломаную» траекторию, мы можем, если захотим, использовать ее в качестве начального приближения для решения исходной задачи с помощью квазилинеаризации. Во многих случаях «ломаная» траектория обеспечивает достаточно точное решение.

7. Динамическое программирование и дифференциальная аппроксимация

Теперь кратко опишем способы, которыми мы можем комбинировать динамическое программирование и дифференциальную аппроксимацию. Предположим, что рассматриваемая нами задача состоит в определении таких параметров a и c, что решение уравнения

$$\frac{dv}{dt} = g(v, a), \quad v(0) = c, \tag{7.1}$$

близко к заданной функции $u(t)$ в смысле минимума выражения

$$\int_0^T (u - v)^2 dt. \tag{7.2}$$

Как и в случае задачи минимизации, обсужденной в предыдущих разделах, метод квазилинеаризации может оказаться несостоятельным, если начальное приближение недостаточно близко к решению.

Одна из причин этой неудачи может состоять в том, что $u(t)$ на самом деле составлена из кусков ряда различных процессов. Таким образом, если мы запишем $v(t) = v(t, a, c)$ (обозначая явно зависимость от a и c), то вполне возможно, что

$$u(t) = \begin{cases} = v(t, a_1, c_1), & 0 \leqslant t \leqslant t_1, \\ = v(t, a_2, c_2), & t_1 \leqslant t \leqslant t_2, \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ = v(t, a_N, c_N), & t_{N-1} \leqslant t \leqslant t_N. \end{cases} \tag{7.3}$$

Как определить величины a_i, c_i и t_i, если мы знаем, что $u(t)$ имеет такую структуру (весьма типичная ситуация во многих случаях)? Это и есть задача идентификации систем.

8. Динамическое программирование и идентификация систем

Чтобы показать, каким образом мы намереваемся применить здесь динамическое программирование, рассмотрим простейший вариант предыдущей задачи, состоящий в нахождении оптимальной аппроксимации на отрезке $[0, T]$ заданной функции $u(t)$ ломаными. Предположим, что допустимы только ломаные, состоящие из $(N+1)$ отрезков прямых, как показано на

рис. 7.2. Мы хотим определить N значений t_i и $2N$ параметров a_i и b_i так, чтобы минимизировать величину

$$D_N = \sum_{i=0}^{N} \int_{t_i}^{t_{i+1}} (u(t) - a_i - b_i t)^2 \, dt. \tag{8.1}$$

Здесь $t_0 = 0$, $t_{N+1} = T$. Введем для любых чисел s_1, s_2, удовлетворяющих условиям $0 \leqslant s_1 \leqslant s_2 \leqslant T$, величину

$$\Delta(s_1, s_2) = \min_{a_i, b_i} \int_{s_1}^{s_2} (u(t) - a_i - b_i t)^2 \, dt. \tag{8.2}$$

Функция $\Delta(s_1, s_2)$ легко получается из решения двух линейных алгебраических уравнений, составленных путем приравнивания нулю частных производных интеграла (8.2) по a_i и b_i.

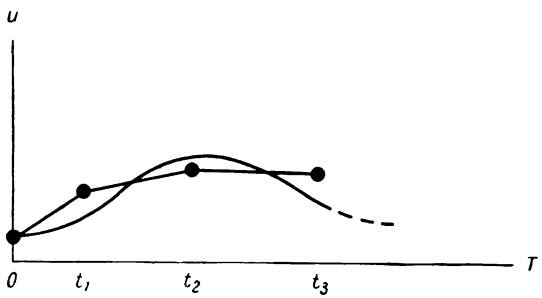


Рис. 7.2.

Введем функции

$$f_N(a) = \min_{a_i, b_i, t_i} D_N, \tag{8.3}$$

определенные для $N = 0, 1, 2, \ldots$, $a \geqslant 0$. Тогда из принципа оптимальности получаем рекуррентное соотношение

$$\begin{aligned} f_N(a) &= \min_{0 \leqslant t_N \leqslant a} [\Delta(t_N, a) + f_{N-1}(t_N)], \quad N \geqslant 1, \\ f_0(a) &= \Delta(0, a). \end{aligned} \tag{8.4}$$

Когда функции $\Delta(t_N, a)$ вычислены, определение последовательности $\{f_N(a)\}$ производится обычным способом.

9. Обсуждение

Величина $\Delta(s_1, s_2)$ возникает в связи с приближением функции $u(t)$ решениями уравнения

$$\frac{d^2v}{dt^2} = 0, \quad v(0) = a, \quad v'(0) = b. \tag{9.1}$$

Это простейшая форма дифференциальной аппроксимации. Если вместо решений уравнения (8.1) для получения аппроксимирующей кривой мы хотим привлечь более сложную дифференциальную аппроксимацию, то следует воспользоваться квазилинеаризацией, как описано в гл. VI. Здесь опять существует ряд методов, которые можно применить для уменьшения времени, требуемого на вычисление функции $\Delta(s_1, s_2)$ для сетки значений (s_1, s_2) на отрезке $0 \leqslant s_1 \leqslant s_2 \leqslant T$.

Наконец отметим, что, никоим образом не увеличивая вычислительные трудности, можно вместо среднеквадратичного критерия использовать критерий Чебышева

$$\min_{a_i, b_i} \max_{s_1 \leqslant t \leqslant s_2} |u(t) - a_i - b_i t|. \tag{9.2}$$

КОММЕНТАРИИ И БИБЛИОГРАФИЯ

1—3. О введении в теорию динамического программирования с приложениями ко многим областям см.

Беллман Р., Динамическое программирование, ИЛ, М., 1960;
Беллман Р., Дрейфус С., Прикладные задачи динамического программирования, изд-во «Наука», М., 1965;
Беллман Р., Процессы регулирования с адаптацией, изд-во «Наука», М., 1964.

Теория положительных операторов со многими дальнейшими ссылками изложена в работах

Walter W., Differential- und Integral-Ungleichungen, Berlin, 1964;
Беккенбах Э., Беллман Р., Неравенства, изд-во «Мир», М., 1965.

4. Эти методы были разработаны в одной совместной работе с С. Азеном.

5. Что касается детального обсуждения транспортных задач, см. цитированную выше книгу Беллмана и Дрейфуса, а также работы

Kalaba R., On some communication network problems, American Mathematical Society, 1960;
Pollack M., Wiebenson N., Solutions of the shortest-route problem: A review, *Operations Res.*, 8 (1960), 224—230;
Bentley D. L., Cooke K. L., Convergence of successive approximations in the shortest-route problem, *J. Math. Anal. Appl.* (в печати).

* Отметим здесь, что эффективные численные алгоритмы решения вариационных задач, связанные с методами динамического программирования, разрабатывались в работах советских авторов. Обзор этих работ см. в статье

Моисеев Н. Н., Численные методы теории оптимальных управлений, использующие вариации в пространстве состояний, *Кибернетика*, № 3 (1966), 1—29.

8. См.

Bellman R., On the approximation of curves by line segments using dynamic programming, *Comm. ACM*, **4** (1961), 284;
Bellman R., Gluss B., Roth R., On the identification of systems and the unscrambling of data: Some problems suggested by neurophysiology, *Proc. Nat. Acad. Sci. USA*, **52** (1964), 1239—1240;
Bellman R., Gluss B., Roth R., Segmental differential approximation and the «Black box» problem, *J. Math. Anal Appl.* (в печати);
Bellman R., Kagiwada H., Kalaba R., Dynamic programming and an inverse problem in neutron transport theory, The RAND Corporation, RM-4495-PR, March 1965.

Приложения

Приложение 1

Программа минимизации времени

```
 260515            TRAJECTORY

C         FOR EULER EQUATION  1.+ (Y')**2 + Y Y'' = 0.
C
       COMMON T,N,C,K
       DIMENSION T(500),C(10)
C
  20  READ 10,XA,YA,XB,YB,H,NHP,NFP,L
  10  FORMAT(5E12.4,3I3)
C
      PRINT 100,XA,YA,XB,YB,H,NHP,NFP,L
 100  FORMAT(1H15X5HINPUT/ 1H05X23HXA,YA,XB,YB,H,NHP,NFP,L /
     1  1H05X 2(1P2E14.4,3X),E14.4, 0P3I6)
C
        NH=(XA-XB)/H
        NN=NH/NHP
C
C            L=HIGHEST APPROXIMATION
C            (XA,YA),  INITIAL POINT
C            (XB,YB),  FINAL POINT
C
      MAX=L+1
      C(1)=(YB-YA)/(XB-XA)
            KM=0
      PRINT 103,KM,C(1)
 103    FORMAT(1H048X,I2,10X 2HC=,1PE14.6)
C
      DO 3 K=2,MAX
            N=2*K+4
C
       DO 30 I=1,500
  30   T(I)=0.0
       T(2)=XA
       T(3)=H
            KM=K-1
   7   DO 1 I=1,KM
            J=2*I+2
       T(J)=YA
   1   T(J+1)=C(I)
   6   T(N-2)=0.
       T(N-1)=0.
       T(N)  =0.
       T(N+1)=1.
       T(N+2)=1.
       T(N+3)=0.
C
      PRINT 104
 104  FORMAT(1H142X1HX,20X4HU(X))
C
      CALL INT(T,N,1,0.,0.,0.,0.,0.,0.)
C
      DO 8 I=1,NN
        DO 9 J=1,NHP
   9    CALL INTM
   8  PRINT 106,T(2),T(N-4)
 106  FORMAT(29X1P2E21.6)
C
```

```
 260515            TRAJECTORY

      C(K)=(YB-T(N-2)-YA*T(N+2))/T(N)
    3 PRINT 103,KM,C(K)
C
C                  FINAL APPROXIMATION AND EXACT SOLUTION
       XI=0.5*(XB+XA+(YB**2-YA**2)/(XB-XA))
       RS=YB**2+(XB-XI)**2
      PRINT 105, XI,RS
 105  FORMAT(1H1,28X,1P2E21.6)
C
           K=MAX+1
           N=2*K+4
C
       DO 31 I=1,500
   31  T(I)=0.0
       T(2)=XA
       T(3)=H
       DO 11 I=1,MAX
         J=2*I+2
       T(J)=YA
   11  T(J+1)=C(I)
       T(N-2)=0.
       T(N-1)=0.
       T(N)=0.
       T(N+1)=1.
       T(N+2)=1.
       T(N+3)=0.
C
      PRINT 101
  101  FORMAT(1H1,42X,1HX,20X,4HU(X),15X,8HEXACT  U)
C
      CALL INT(T,N,1,0.,0.,0.,0.,0.,0.)
C
        X=XA
        F=NFP
        D=F*H
        NNN=NH/NHP
      DO 2 I=1,NNN
        DO 19 J=1,NFP
   19 CALL INTM
       X=X+D
       Y=SQRTF(ABSF(RS-(X-XI)**2))
    2 PRINT 102,T(2),T(N-4),Y
  102  FORMAT(29X,1P3E21.6)
C
      C(K)=(YB-T(N-2)-YA*T(N+2))/T(N)
      CA=(XI-XA)/YA
      PRINT 103,MAX,C(K)
      PRINT 107,CA
  107 FORMAT(1H0,60X,2HC=,1PE14.6)
C
      GO TO 20
      END(1,1,0,0,0,0,1,0,0,0,0,0,0,0,0)
```

```
260513                          DAUX

      SUBROUTINE DAUX
       COMMOM T,N,C,K
       DIMENSION T(500),C(10)
C
       T(N+4)=C(1)
       T(N+5)=0.0
C
      DO 1 I=2,K
            J=2*I+2
            M=N+J
       T(M)=T(J+1)
            AA=-2.*T(J-1)/T(J-2)
            BB=(1.+T(J-1)**2)/T(J-2)**2
            CC=-2./T(J-2)
    1  T(M+1)=AA*T(J+1)+BB*T(J)+CC
       T(M+2)=T(J+3)
       T(M+3)=AA*T(J+3)+BB*T(J+2)
       T(M+4)=T(J+5)
       T(M+5)=AA*T(J+5)+BB*T(J+4)
         RETURN
      END(1,1,0,0,0,0,0,1,0,0,0,0,0,0,0,0)*
```

Приложение 2

Программа для задач управления и проектирования

```
260550              DESIGN PROBLEM

      COMMON N1,KMAX,NP,NMAX,HGRID,CX,T,X,Y,U,V,W,PREV,H,P,A,B,C,NEQ
      DIMENSION T(250),W(4,500),H(4,4,500),P(4,500),A(2,2),B(2),C(4),
     1  PREV(4)
C             X,  STATE VARIABLE
C             Y,  CONTROL VARIABLE
C             U(=MU),  LAGRANGE MULTIPLIER
C             V(=A),  DESIGN PARAMETER
C
    1 CALL INPUT
    2 CALL START
C
C                    K ITERATIONS
C
    3 DO 19 K=1,KMAX
C                  INITIALIZE
      DO 5 I=1,250
    5 T(I)=0.
      T(2)=0.
      T(3)=HGRID
      T(4)=1.
      T(9)=1.
      T(14)=1.
      T(19)=1.
      N=0
      PRINT 111, (T(L),L=4,43)
      NEQ=20
      CALL INT(T,NEQ,N1,0.,0.,0.,0.,0.,0.)
C             INTEGRATE P'S AND H'S, STORING GRID VALUES
      N=0
    8 DO 11 M=1,NP
      N=N+1
        X=W(1,N)
        Y=W(2,N)
        U=W(3,N)
        V=W(4,N)
        CALL INTM
        L=3
        DO 9 I=1,4
          DO 9 J=1,4
          L=L+1
    9   H(I,J,N)=T(L)
        DO 10 J=1,4
          L=L+1
   10   P(J,N)=T(L)
   11 CONTINUE
C
    6 PRINT 111,N,(T(L),L=4,43)
      PRINT 111,N,((H(I,J,N),J=1,4),I=1,4),(P(J,N),J=1,4)
  111 FORMAT(1H019X110,4E20.6/(30X4E20.6))
      IF(N-NMAX)8,7,7
    7 PRINT 40,K
   40   FORMAT(1H0/65X9HITERATION,I3)
C
C                  EVALUATE CONSTANTS
C
```

```
 260550            DESIGN PROBLEM

      N=NMAX
      A(1,1)=H(2,2,N)
      A(2,1)=H(2,3,N)
      A(1,2)=H(3,2,N) + H(4,2,N)
      A(2,2)=H(3,3,N) + H(4,3,N)
C
      B(1)=-P(2,N) - CX*H(1,2,N)
      B(2)=-P(3,N) - CX*H(1,3,N)
C
      DET=A(1,1)*A(2,2) - A(2,1)*A(1,2)
      G=ABSF(DET)
      IF(G-.000001)200,200,12
C
  200 PRINT 201,DET,((A(I,J),J=1,2),B(I),I=1,2)
  201   FORMAT(1H020X12HDETERMINANT=,E16.6/(40X3E20.6))
      CALL EXIT
C
   12 C(1)=CX
      C(2)=(B(1)*A(2,2) - B(2)*A(1,2))/DET
      C(3)=(B(2)*A(1,1) - B(1)*A(2,1))/DET
      C(4)=C(3)
      GRID=0.
      PRINT 45, GRID,(C(I),I=1,4)
   45   FORMAT(1H026X4HGRID,14X1HX,19X1HY,18X2HMU,19X1HA/
     1  20XF10.4,4E20.6)
C
C                   NEW VALUES OF VARIABLES
      N=0
   13 DO 14 M=1,NP
        N=N+1
        DO 14 I=1,4
        W(I,N)=P(I,N)
        DO 14 J=1,4
   14   W(I,N)=W(I,N) + C(J)*H(J,I,N)
      FN=N
      GRID=FN*HGRID
      PRINT 50, GRID,(W(I,N),I=1,4)
      IF(N-NMAX)13,15,15
C
C                   COMPARE C'S, PRESENT AND PREVIOUS
   15 DO 16 I=1,4
        G=ABSF(C(I)-PREV(I))
        IF(G-.000001)16,16,17
   16 CONTINUE
           GO TO 1
   17 DO 18 I=1,4
   18 PREV(I)=C(I)
C
   19 CONTINUE
           GO TO 1
C
   50   FORMAT(20XF10.4,4E20.6)
      END(1,1,0,0,0,0,1,0,0,0,0,0,0,0,0)
```

```
260552            INPUT - DESIGN

      SUBROUTINE INPUT
       COMMON N1,KMAX,NP,NMAX,HGRID,CX,T,X,Y,U,V,W,PREV,H,P,A,B,C
       DIMENSION T(250),W(4,500),H(4,4,500),P(4,500),A(2,2),B(2),C(4),
     1  PREV(4)
C            N1, INTEGRATION OPTION
C            KMAX, NUMBER OF ITERATIONS
C            NP, NUMBER OF GRIDS BEFORE PRINT-OUT
C            NMAX, TOTAL NUMBER OF INTEGRATIONS IN INTERVAL
C            HGRID, GRID SIZE
C            CX, X(0)
C
       READ 110,N1,KMAX,NP,NMAX,HGRID,CX
       PRINT 10,N1,KMAX,NP,NMAX,HGRID,CX
  110    FORMAT(4I4,2E8.2)
   10    FORMAT(1H1 20X14HDESIGN PROBLEM//
     1  38X2HN1,6X4HKMAX,8X2HNP,6X4HNMAX,13X5HHGRID,16X4HX(0)/
     2  30X4I10,2E20.6)
         RETURN
       END(1,1,0,0,0,0,1,0,0,0,0,0,0,0,0)
```

```
260551            START  (FOR 260550)

      SUBROUTINE START
       COMMON N1,KMAX,NP,NMAX,HGRID,CX,T,X,Y,U,V,W,PREV,H,P,A,B,C
       DIMENSION T(250),W(4,500),H(4,4,500),P(4,500),A(2,2),B(2),C(4),
     1  PREV(4)
C            INITIALLY T=0. X(T)=CX, Y(T)=0. MU(T)=0. A(T)=0
      DO 1 N=1,NMAX
      W(1,N)=CX
      W(2,N)=0.
      W(3,N)=0.
    1 W(4,N)=0.
C
      DO 2 I=1,4
    2 PREV(I)=W(I,1)
C
      K=0
      PRINT 40,K
      GRID=0.
      PRINT 45,GRID,(W(I,1),I=1,4)
   40   FORMAT(1H0/65X9HITERATION,I3)
   45   FORMAT(1H026X4HGRID,14X1HX,19X1HY,18X2HMU,19X1HA/20XF10.4,4E20.6
     1  /1H019X22HFOR ALL VALUES OF TIME)
         RETURN
      END(1,1,0,0,0,0,1,0,0,0,0,0,0,0,0)
```

```
260553              DAUX  (FOR DESIGN)

      SUBROUTINE DAUX
       COMMON N1,KMAX,NP,NMAX,HGRID,CX,T,X,Y,U,V,W,PREV,H,P,A,B,C,NEG
       DIMENSION T(250),W(4,500),H(4,4,500),P(4,500),A(2,2),B(2),C(4),
     1  PREV(4)
C
      L=20
      DO 1 I=1,4
        L=L+4
      T(L)=-V*T(L-20) + T(L-19) -X*T(L-17)
      T(L+1)=T(L-20) + V*T(L-19) + Y*T(L-17)
      T(L+2)=Y*T(L-20) + X*T(L-19)
      T(L+3)=0.
C
      T(40)=-V*T(20) + T(21) - X*T(23) + V*X
      T(41)=T(20) + V*T(21) + Y*T(23) - V*Y
      T(42)=Y*T(20) + X*T(21) - X*Y
      T(43)=0.
        RETURN
      END(1,1,0,0,0,0,1,0,0,0,0,0,0,0,0)
```

Программа для обратной задачи переноса излучения (два слоя, три константы)

```
$JOB            2609,STRAT3,HK0160,5,0,20,P
$PAUSE
$IBJOB STRAT2  MAP
$IBFTC RTINV
      COMMON N,RT(7),WT(7),WR(7),AR(7,7),NPRNT,M1MAX,KMAX,DELTA,XTAU,
     1  XLAM(3),        B2(7,7),R2(7,7),IFLAG,R(28,101),T(1491),SIG,
     2  P(28,101),H(28,3,101),PLAM(3),HLAM(3,3),P2(7,7),
     3  H2(7,7,3),CONST(3),NEQ
C
C            PHASE I
C
    1  READ1000,N
       PRINT899
       PRINT900,N
       READ1001,(RT(I),I=1,N)
       PRINT901,(RT(I),I=1,N)
       READ1001,(WT(I),I=1,N)
       PRINT901,(WT(I),I=1,N)
       DO 2 I=1,N
       WR(I)=WT(I)/RT(I)
       DO 2 J=1,N
    2  AR(I,J)= 1.0/RT(I) + 1.0/RT(J)
C
  899 FORMAT(1H1,46X,36HRADIATIVE TRANSFER - INVERSE PROBLEM / )
 1000 FORMAT(6I12)
  900 FORMAT(6I20)
 1001 FORMAT(6E12.8)
  901 FORMAT(6E20.8)
       READ1000,NPRNT,M1MAX,KMAX
       PRINT900,NPRNT,M1MAX,KMAX
       READ1001,DELTA
       PRINT901,DELTA
       READ1001,XTAU,(XLAM(I),I=1,3)
       PRINT902
       PRINT903,XTAU,(XLAM(I),I=1,3)
  902 FORMAT(1H1,23HPHASE I - TRUE SOLUTION /)
  903 FORMAT(1HO/
      1     1X,11HTHICKNESS =, F10.4 /
      2     1X,11HALBEDO(X) =, 20HA + B*TANH(10*(X-C))  //
      3  1X,3HA =, E16.8, 10X,3HB =, E16.8, 10X,3HC =, E16.8 //)
       CALL NONLIN
       DO 3 I=1,N
       DO 3 J=1,N
    3  B2(I,J)=R2(I,J)
C
C
C            PHASE II
C
    4  READ1001,XTAU,(XLAM(I),I=1,3)
       K=0
       PRINT904,K
       PRINT903,XTAU,(XLAM(I),I=1,3)
C
       CALL NONLIN
```

```
C
  904 FORMAT(1H1   13HAPPROXIMATION, I3/ )
C
C            QUASILINEARIZATION ITERATIONS
C
      DO 5 K1=1,KMAX
      PRINT904,K1
      CALL PANDH
      CALL LINEAR
    5 CONTINUE
C
C
C
      READ1000,IGO
      GO TO (1,4),IGO
      END
$IBFTC DAUX    LIST
      SUBROUTINE DAUX
      DIMENSION V2(7,7),X(3),F(7),G(7)
     1  ,VLAM(3)
      COMMON N,RT(7),WT(7),WR(7),AR(7,7),NPRNT,M1MAX,KMAX,DELTA,XTAU,
     1  XLAM(3),       B2(7,7),R2(7,7),IFLAG,R(28,101),T(1491),SIG,
     2  P(28,101),H(28,3,101),PLAM(3),HLAM(3,3),P2(7,7),
     3  H2(7,7,3),CONST(3),NEQ
      GO TO (1,2),IFLAG
C
CNONLINEAR
C
    1 L=3
      DO 4 I=1,N
      DO 4 J=1,I
      L=L+1
    4 V2(I,J)=T(L)
      DO 5 I=1,N
      DO 5 J=I,N
    5 V2(I,J)=V2(J,I)
      DO 51 I=1,3
      L=L+1
   51 VLAM(I)=T(L)
      SIG=T(2)
      Y=XTAU*SIG
      DO 52 I=1,3
   52 X(I)=VLAM(I)
      CALL ALBEDO(Y,X,Z)
      ZLAMDA=Z
C
      DO 6 I=1,N
      F(I)=0.0
      DO 7 K=1,N
    7 F(I)=F(I) + WR(K)*V2(I,K)
    6 F(I)=0.5*F(I) + 1.0
C
      DO 8 I=1,N
      DO 8 J=1,I
```

```
      L=L+1
      DR=-AR(I,J)*V2(I,J) + ZLAMDA*F(I)*F(J)
    8 T(L)=DR
      DO 9 I=1,3
      L=L+1
    9 T(L)=0.0
      RETURN
C
C
CLINEAR
C
    2 SIG=T(2)
      Y=XTAU*SIG
      DO 21 I=1,3
   21 X(I)=XLAM(I)
      CALL ALBEDO(Y,X,Z)
      ZLAMDA=Z
C
      DO 16 I=1,N
      F(I)=0.0
      DO 17 K=1,N
   17 F(I)=F(I) + WR(K)*R2(I,K)
   16 F(I)=0.5*F(I) + 1.0
C
CP'S
C
      L=3
      DO 14 I=1,N
      DO 14 J=1,I
      L=L+1
   14 V2(I,J)=T(L)
      DO 15 I=1,N
      DO 15 J=I,N
   15 V2(I,J)=V2(J,I)
      DO 18 I=1,3
      L=L+1
   18 VLAM(I)=T(L)
C
      DO 10 I=1,N
      G(I)=0.0
      DO 10 K=1,N
   10 G(I)=G(I) + (V2(I,K)-R2(I,K))*WR(K)
      ARG=10.D*(Y-XLAM(3))
      TARG=TANH(ARG)
      XTANX=-10.0*XLAM(2)*(1.0-TARG**2)
        M=3+NEQ
      DO 12 I=1,N
      DO 12 J=1,I
      FIJ=F(I)*F(J)
      CAPF=-AR(I,J)*R2(I,J) + ZLAMDA*FIJ
      T1=CAPF
      T2=-AR(I,J)*(V2(I,J)-R2(I,J))
     1  + 0.5*ZLAMDA*(F(I)*G(J) + F(J)*G(I))
      T3=(VLAM(1)-XLAM(1))*FIJ
```

```
      T4=(VLAM(2)-XLAM(2))*TARG*FIJ
      T5=(VLAM(3)-XLAM(3))*XTANX*FIJ
        M=M+1
   12 T(M)=T1+T2+T3+T4+T5
      DO 19 I=1,3
        M=M+1
   19 T(M)=0.0
C
CH'S
C
      DO 100 K=1,3
C
      DO 24 I=1,N
      DO 24 J=1,I
      L=L+1
   24 V2(I,J)=T(L)
      DO 25 I=1,N
      DO 25 J=I,N
   25 V2(I,J)=V2(J,I)
      DO 26 I=1,3
      L=L+1
   26 VLAM(I)=T(L)
C
      DO 20 I=1,N
      G(I)=0.0
      DO 20 J=1,N
   20 G(I)=G(I) +  V2(I,J)*WR(J)
C
      DO 22 I=1,N
      DO 22 J=1,I
      FIJ=F(I)*F(J)
      T1=0.0
      T2=-AR(I,J)*V2(I,J) + 0.5*ZLAMDA*(F(I)*G(J) + F(J)*G(I))
      T3=VLAM(1)*FIJ
      T4=VLAM(2)*TARG*FIJ
      T5=VLAM(3)*XTANX*FIJ
        M=M+1
   22 T(M)=T1+T2+T3+T4+T5
C
      DO 29 I=1,3
        M=M+1
   29 T(M)=0.0
  100 CONTINUE
      RETURN
      END
$IBFTC NONLIN
      SUBROUTINE NONLIN
      COMMON N,RT(7),WT(7),WR(7),AR(7,7),NPPN1,M1MAX,KMAX,DELTA,XTAU,
     1  XLAM(3),        B2(7,7),R2(7,7),IFLAG,P(28,101),T(1491),SIG,
     2  P(28,101),H(28,3,101),PLAM(3),HLAM(3,3),P2(7,7),
     3  H2(7,7,3),CONST(3),NEQ
C           NONLINEAR D.E. FOR TRUE SOLUTION OR FOR INITIAL APPROX.
C
      IFLAG=1
```

```
      T(2)=0.0
      T(3)=DELTA
      M=1
      L1=0
      L3=3
      DO 1 I=1,N
      DO 1 J=1,I
      L1=L1+1
      L3=L3+1
      R2(I,J)=0.0
      R(L1,M)=R2(I,J)
    1 T(L3)=R2(I,J)
      DO 2 I=1,3
      L3=L3+1
    2 T(L3)=XLAM(I)
C
      NEQ=(N*(N+1))/2 + 3
      CALL INTS(T,NEQ,2,0,0,0,0,0,0)
C
      SIG=T(2)
      CALL OUTPUT
C
      DO 5 M1=1,M1MAX
      DO 4 M2=1,NPRNT
      CALL INTM
      M=M+1
      L1=0
      L3=3
      DO 3 I=1,N
      DO 3 J=1,I
      L1=L1+1
      L3=L3+1
      R2(I,J)=T(L3)
    3 R(L1,M)=R2(I,J)
    4 SIG=T(2)
    5 CALL OUTPUT
C
      RETURN
      END
$IBFTC PANDH
      SUBROUTINE PANDH
      COMMON N,RT(7),WT(7),WR(7),AR(7,7),NPRNT,M1MAX,KMAX,DELTA,XTAU
     1  XLAM(3),        B2(7,7),R2(7,7),IFLAG,R(28,101),T(1491),SIG,
     2  P(28,101),H(28,3,101),PLAM(3),HLAM(3,3),P2(7,7),
     3  H2(7,7,3),CONST(3),NEQ
      IFLAG=2
      T(2)=0.0
      T(3)=DELTA
      M=1
C P=S
C
      L1=0
      L3=3
      DO 1 I=1,N
```

```
      DO 1 J=1,I
      L1=L1+1
      L3=L3+1
      P(L1,M)=0.0
    1 T(L3)=P(L1,M)
      DO 2 I=1,3
      L3=L3+1
      PLAM(I)=0.0
    2 T(L3)=PLAM(I)
C
C H S
C
      DO 7 K=1,3
      L1=0
      DO 3 I=1,N
      DO 3 J=1,I
      L1=L1+1
      L3=L3+1
      H(L1,K,M)=0.0
    3 T(L3)=H(L1,K,M)
C
      DO 7 I=1,3
      L3=L3+1
      HLAM(I,K)=0.0
      IF(I-K)7,6,7
    6 HLAM(I,K)=1.0
    7 T(L3)=HLAM(I,K)
C
      L=0
      DO 8 I=1,N
      DO 8 J=1,I
      L=L+1
    8 R2(I,J)=R(L,M)
      DO 9 I=1,N
      DO 9 J=I,N
    9 R2(I,J)=R2(J,I)
C
      NEQ=4*((N*(N+1))/2 + 3)
      CALL INTS(T,NEQ,2.0,0.0,0.0,0.0,0.0)
      LMAX=(N*(N+1))/2
      PRINT52,T(2),(P(L,M),H(L,1,M),L=1,LMAX)
   52 FORMAT(1H0F9.4,5E20.8/(10X5E20.8))
C
      DO 51 M1=1,M1MAX
      DO 50 M2=1,NPRNT
      CALL INTM
      M=M+1
CPREV.APPROX. R(I,J)
      L1=0
      DO 10 I=1,N
      DO 10 J=1,I
      L1=L1+1
   10 R2(I,J)=R(L1,M)
      DO 11 I=1,N
```

```
      DO 11 J=I,N
   11 R2(I,J)=R2(J,I)
      L1=0
      L3=3
      DO 12 I=1,N
      DO 12 J=1,I
      L1=L1+1
      L3=L3+1
   12 P(L1,M)=T(L3)
      L3=L3+3
      DO 13 K=1,3
      L1=0
      DO 14 I=1,N
      DO 14 J=1,I
      L1=L1+1
      L3=L3+1
   14 H(L1,K,M)=T(L3)
   13 L3=L3+3
   50 CONTINUE
   51 PRINT52,T(2),(P(L,M),H(L,1,M),L=1,LMAX)
      RETURN
      END
$IBFTC LINEAR
      SUBROUTINE LINEAR
      DIMENSION CHK1(3)
      DIMENSION A(49,3),B(49),EMAT(50,50),          PIVOT(50),INDEX(50,2)
     1),IPIVOT(50),FVEC(50,1)
      COMMON N,RT(7),WT(7),WR(7),AR(7,7),NPRNT,M1MAX,KMAX,DELTA,XTAU,
     1  XLAM(3),        B2(7,7),R2(7,7),IFLAG,R(28,101),T(1491),SIG,
     2  P(28,101),H(28,3,101),PLAM(3),HLAM(3,3),P2(7,7),
     3  H2(7,7,3),CONST(3),NEQ
CBOUNDARY CONDITIONS
      MLAST=NPRNT*M1MAX + 1
      DO 1 K=1,3
      L=0
      DO 2 I=1,N
      DO 2 J=1,I
      L=L+1
   2  H2(I,J,K)=H(L,K,MLAST)
      DO 1 I=1,N
      DO 1 J=I,N
   1  H2(I,J,K)=H2(J,I,K)
      L=0
      DO 3 I=1,N
      DO 3 J=1,I
      L=L+1
   3  P2(I,J)=P(L,MLAST)
      DO 4 I=1,N
      DO 4 J=I,N
   4  P2(I,J)=P2(J,I)
CLEAST SQUARES
      DO 5 K=1,3
      L=0
      DO 5 I=1,N
```

```
      DO 5 J=1,N
      L=L+1
    5 A(L,K)=H2(I,J,K)
      L=0
      DO 6 I=1,N
      DO 6 J=1,N
      L=L+1
    6 B(L)=B2(I,J) - P2(I,J)
C
      LMAX=N**2
      PRINT60
   60 FORMAT(1H0)
      DO 61 L=1,LMAX
   61 PRINT82,(A(L,K),K=1,3),B(L)
C
      DO 8 I=1,3
      DO 7 J=1,3
      SUM=0.0
      DO 9 L=1,LMAX
    9 SUM=SUM + A(L,I)*A(L,J)
    7 EMAT(I,J)=SUM
      SUM=0.0
      DO 10 L=1,LMAX
   10 SUM=SUM + A(L,I)*B(L)
    8 FVEC(I,1)=SUM
C
      PRINT60
      DO 81 I=1,3
   81 PRINT82,(EMAT(I,J),J=1,3),FVEC(I,1)
   82 FORMAT(10X6E20.8)
C
      CALL MATINV(EMAT,3,FVEC,1,DETERM,PIVOT,INDEX,IPIVOT)
C
      DO 11 I=1,3
   11 CONST(I)=FVEC(I,1)
C
      DO 20 I=1,3
   20 XLAM(I)=CONST(I)
      PRINT903,XTAU,(XLAM(I),I=1,3)
  903 FORMAT(1H0/
     1  1X11HTHICKNESS =, E16.8 /
     2     1X11HALBEDO(X) =, 20HA + B*TANH(10*(X-C))  //
     3  1X3HA =, E16.8, 10X3HB =, E16.8, 10X3HC =, E16.8 //)
C
CNEW APPROXIMATION
C
      M=1
      L=0
      DO 12 I=1,N
      DO 12 J=1,I
      L=L+1
      SUM=P(L,M)
      DO 13 K=1,3
   13 SUM =SUM + CONST(K)*H(L,K,M)
```

```
   12  R(L,M)=SUM
       L=0
       DO 14 I=1,N
       DO 14 J=1,I
       L=L+1
   14  R2(I,J)=R(L,M)
       SIG=0.0
       CALL OUTPUT
C
       DO 50 M1=1,M1MAX
       DO 18 M2=1,NPRNT
       M=M+1
       L=0
       DO 15 I=1,N
       DO 15 J=1,I
       L=L+1
       SUM=P(L,M)
       DO 16 K=1,3
   16  SUM=SUM + CONST(K)*H(L,K,M)
   15  R(L,M)=SUM
       L=0
       DO 17 I=1,N
       DO 17 J=1,I
       L=L+1
   17  R2(I,J)=R(L,M)
   18  SIG=SIG + DELTA
   50  CALL OUTPUT
C
       RETURN
       END
$IBFTC OUTPUT
       SUBROUTINE OUTPUT
       DIMENSION X(3)
       COMMON N,RT(7),WT(7),WR(7),AR(7,7),NPRNT,M1MAX,KMAX,DELTA,XTAU,
      1  XLAM(3),        B2(7,7),R2(7,7),IFLAG,R(28,101),T(149),SIG,
      2  P(28,101),H(28,3,101),PLAM(3),HLAM(3,3),P2(7,7),
      3  H2(7,7,3),CONST(3),NEQ
       DO 1 I=1,N
       DO 1 J=I,N
    1  R2(I,J)=R2(J,I)
       Y=XTAU*SIG
       DO 3 I=1,3
    3  X(I)=XLAM(I)
       CALL ALBEDO(Y,X,Z)
       PRINT100, SIG,Y,Z
  100  FORMAT(1H0 7HSIGMA =,F6.2, 4X5HTAU =, F6.2, 4X8HALBEDO =,F6.2/)
       DO 2 J=1,N
     2 PRINT101,J,(R2(I,J),I=1,N)
  101  FORMAT(I10, 7F10.6)
       RETURN
       END
$IBFTC ALBEDO
       SUBROUTINE ALBEDO(Y,X,Z)
       DIMENSION X(3)
```

```
      COMMON N,RT(7),WT(7),WR(7),AR(7,7),NPRNT,M1MAX,KMAX,DELTA,XTAU,
     1  XLAM(3),       B2(7,7),R2(7,7),IFLAG,R(28,101),T(1491),SIG,
     2  P(28,101),H(28,3,101),PLAM(3),HLAM(3,3),P2(7,7),
     3  H2(7,7,3),CONST(3),NEQ
      ARG=10.0*(Y-X(3))
      Z=X(1) + X(2)*TANH(ARG)
      RETURN
      END
$ENTRY          RTINV
           7
25446046E-0112923441E-0029707742E-0050000000E 0070292258E 0087076559E 00   032
97455396E 00                                                               033
64742484E-0113985269E-0019091502E-0020897958E-0019091502E-0013985269E-00   0032
64742484E-01                                                               0033
           10           10            4
         0.01
          1.0          0.5          0.1          0.5
          1.0          0.6          .09          0.4
                                                                           504

                                                                           504
```

Программа для уравнения Ван-дер-Поля

```
CVDP021              VAN DER POL - PHASES II, III
      COMMON SCRACH,T,N1,NMAX,KMAX,HGRID,NTIME,XOBS,W,H,P,A,B,X,U,Z,TI,
     1          PREV,NOBS,IFLAG
      DIMENSION SCRACH(64),T(150),NTIME(100),XOBS(100),W(3,1001),
     1  H(3,3,1001),P(3,1001),A(16,16),B(16,1),PREV(3)
C
C           INPUT  NOBS  OBSERVATIONS OF  x  AT KNOWN TIMES
C           DETERMINE SYSTEM PARAMETER LAMBDA (Z IN FORTRAN)
C                 INPUT AND START
    1 CALL INSTRI
C
C                 K ITERATIONS
C
      DO 99 K=1,KMAX
C
      DO 2 I=1,150
    2 T(I)=0.0
      T(2)=TI
      T(3)=HGRID
      T(4)=1.0
      T(8)=1.0
      T(12)=1.0
        X=PREV(1)
        U=PREV(2)
        Z=PREV(3)
      CALL INT(T,12,N1,0.,0.,0.,0.,0.,0.)
        N=1
        L=3
      DO 21 I=1,3
      DO 21 J=1,3
        L=L+1
   21 H(I,J,N)=T(L)
      DO 22 I=1,3
        L=L+1
   22 P(I,N)=T(L)
C
C                 INTEGRATE P'S AND H'S
      DO 4 N=2,NMAX
        X=W(1,N)
        U=W(2,N)
        Z=W(3,N)
        CALL INTM
          L=3
        DO 3 I=1,3
        DO 3 J=1,3
          L=L+1
    3   H(I,J,N)=T(L)
        DO 4 I=1,3
          L=L+1
    4   P(I,N)=T(L)
C
C                 DETERMINE CONSTANTS, OR INITIAL VALUES
      CALL CNSTNT
      TIME=TI
      PRINT 50, K,TIME,(B(I,1),I=1,3)
C
C                 NEW VARIABLES
      DO 7 N=2,NMAX
        DO 6 I=1,3
```

```
        W(I,N)=P(I,N)
        DO 6 J=1,3
    6   W(I,N)=W(I,N)+B(J,1)*H(J,I,N)
        FN=N-1
        TIME=FN*HGRID+TI
    7 PRINT 70, TIME,(W(I,N),I=1,3)
C
C                   COMPARE CONSTANTS
      DO 8 I=1,3
        G=ABSF(B(I,1)-PREV(I))
        IF(G-.000001)8,8,9
    8 CONTINUE
        GO TO 1
    9 DO 10 I=1,3
   10 PREV(I)=B(I,1)
C
   99 CONTINUE
        GO TO 1
C
   50 FORMAT(1H0//59X9HITERATION,I3//38X1HT,13X1HX,19X1HU,17X6HSYSTEM//
     1  30XF10.2,3E20.6)
   70 FORMAT(30XF10.2,3E20.6)
         END


CVDP022             INPUT-START, VAN DER POL PHASE II,III
      SUBROUTINE  INSTRT
        COMMON SCRACH,T,N1,NMAX,KMAX,HGRID,NTIME,XOBS,W,H,P,A,B,X,U,Z,TI,
     1          PREV,NOBS,IFLAG
        DIMENSION SCRACH(64),T(150),NTIME(100),XOBS(100),W(3,1001),
     1  H(3,3,1001),P(3,1001),A(16,16),B(16,1),PREV(3)
C                   INPUT
      READ 110,N1,NOBS,KMAX,NMAX,HGRID,(NTIME(I),XOBS(I),I=1,NOBS)
        IF(NMAX)9,9,1
    1 PRINT 10,N1,NOBS,KMAX,NMAX,HGRID,(NTIME(I),XOBS(I),I=1,NOBS)
C
C                   START
        IFLAG=1
      READ 120,TI,X,U,Z
        DO 2 I=1,150
    2   T(I)=0.0
        T(2)=TI
        T(3)=HGRID
        T(4)=X
        T(5)=U
        T(6)=Z
      CALL INT(T,3,N1,0.,0.,0.,0.,0.,0.)
        K=0
      PRINT 50, K,T(2),(T(I),I=4,6)
C
      DO 4 N=2,NMAX
        CALL INTM
        DO 3 I=1,3
    3   W(I,N)=T(I+3)
    4 PRINT 70, T(2),(T(I),I=4,6)
C
      IFLAG=2
      PREV(1)=X
      PREV(2)=U
      PREV(3)=Z
```

```
C
        RETURN
    9   CALL EXIT
   10   FORMAT(1H1,30X
     1    58HVAN DER POL - PHASE II - DETERMINATION OF SYSTEM PARAMETER//
     2    20X,4I10,F10.4/(18X,I2,E16.6,I2,E16.6,I2,E16.6))
   50   FORMAT(1H0//59X,9HITERATION,I3//38X,1HT,13X,1HX,19X,1HU,17X,6HSYSTEM//
     1    30XF10.2,3E20.6)
   70   FORMAT(30XF10.2,3E20.6)
  110   FORMAT(4I5,F10.2,2(I4,E11.6)/(4(I4,E11.6)))
  120   FORMAT(4E12.6)
        END


CVDP023             DAUX - VAN DER POL - PHASE II,III
        SUBROUTINE DAUX
        COMMON SCRACH,T,N1,NMAX,KMAX,HGRID,NTIME,XOBS,W,H,P,A,B,X,U,Z,TI,
     1             PREV,NOBS,IFLAG
        DIMENSION SCRACH(64),T(150),NTIME(100),XOBS(100),W(3,1001),
     1    H(3,3,1001),P(3,1001),A(16,16),B(16,1),PREV(3)
C
        GO TO (1,2),IFLAG
C                   NONLINEAR EQUATIONS
    1   T(7)=T(5)
        T(8)=-(T(4)**2-1.)*T(5)*T(6)-T(4)
        T(9)=0.
        RETURN
C                   LINEAR EQUATIONS
    2   XX=X**2
        AA=-2.*X*U*Z-1.
        BB=Z*(1.-XX)
        CC=U*(1.-XX)
        L=13
        DO 3 I=1,4
        L=L+3
        T(L)=T(L-11)
        T(L+1)=AA*T(L-12)+BB*T(L-11)+CC*T(L-10)
    3   T(L+2)=0.0
        T(L+1)=T(L+1)+U*Z*(3.*XX-1.)
        RETURN
        END


CVDP024             CNSTNT - PHASE II
        SUBROUTINE CNSTNT
        COMMON SCRACH,T,N1,NMAX,KMAX,HGRID,NTIME,XOBS,W,H,P,A,B,X,U,Z,TI,
     1             PREV,NOBS,IFLAG
        DIMENSION SCRACH(64),T(150),NTIME(100),XOBS(100),W(3,1001),
     1    H(3,3,1001),P(3,1001),A(16,16),B(16,1),PREV(3)
C
        DO 1 I=1,3
        N=NTIME(I)
        B(I,1)=XOBS(I)-P(1,N)
        DO 1 J=1,3
    1   A(I,J)=H(J,1,N)
        CALL MATINV(A,3,B,1,DET)
        DO 2 I=1,3
    2   W(I,1)=B(I,1)
        RETURN
        END
```

Приложение 5

Программа для определения орбиты

```
C260582              CELESTIAL MECHANICS - PHASE II
      COMMON SCRACH,T,NEQ,N1,NMAX,KMAX,HGRID,A2,A3,A4,A5,A6,A7,
     1  NGRID,THETA,W,H,P,A,B,X,U,Y,V
      DIMENSION SCRACH(64),T(250),NGRID(4),THETA(4),W(4,500),H(4,4,500)
     1  ,P(4,500),A(16,16),B(16,1)
C
    1 CALL INPUT
    2 CALL START(W)
C                    K ITERATIONS
    3 DO 19 K=1,KMAX
        NEQ=20
    4   T(2)=TZERO
        T(3)=HGRID
        DO 5 I=4,43
    5   T(I)=0.
        T(4)=1.
        T(9)=1.
        T(14)=1.
        T(19)=1.
    6   CALL INT(T,NEQ,N1,A2,A3,A4,A5,A6,A7)
C                 INTEGRATE OVER RANGE
    7   DO 11 N=1,NMAX
          X=W(1,N)
          U=W(2,N)
          Y=W(3,N)
          V=W(4,N)
          CALL INTM
C                   STORE P'S AND H'S
          L=3
    8     DO 9 I=1,4
            DO 9 J=1,4
            L=L+1
    9     H(I,J,N)=T(L)
          DO 10 J=1,4
            L=L+1
   10     P(J,N)=T(L)
   11   CONTINUE
C
C                 COMPUTE CONSTANTS
   12   DO 14 I=1,4
          N=NGRID(I)
          THET=THETA(I)
          STHET=SINF(THET)
          CTHET=COSF(THET)
          DO 13 J=1,4
   13     A(I,J)=H(J,1,N)*STHET -H(J,3,N)*CTHET
          B(I,1)=(1.-P(1,N))*STHET +P(3,N)*CTHET
  114     FORMAT(1H0,9X,5E20.6)
   14   PRINT 114,(A(I,JJ),JJ=1,4),B(I,1)
   15   CALL MATINV(A,4,B,1,DET)
        PRINT115,(B(I,1),I=1,4)
  115     FORMAT(1H0,9X,4E2 .6)
C
C                 COMPUTE NEW W'S
        PRINT40,K
   40     FORMAT(1H0/65X 9HITERATION,I3//26X,4HGRID,14X,1HX,19X,2HX',18X,1HY,
     1  19X,2HY')
   16   DO 18 N=1,NMAX
          DO 17 I=1,4
          W(I,N)=P(I,N)
```

```
        DO 17 J=1,4
  17    W(I,N)=W(I,N) + B(J,1)*H(J,I,N)
  50    FORMAT(20X,I10,4E20.6)
  18  PRINT50,N,(W(I,N),I=1,4)
C
  19  CONTINUE
C
       GO TO 1
         END
C260585              INPU1
      SUBROUTINE INPUT
       COMMON SCRACH,T,NEQ,N1,NMAX,KMAX,HGRID,A2,A3,A4,A5,A6,A7,
     1  NGRID,THETA,W,H,P,A,B,X,U,Y,V
       DIMENSION SCRACH(64),T(250),NGRID(4),THETA(4),W(4,500),H(4,4,500)
     1  ,P(4,500),A(16,16),B(16,1)
C
       READ110,N1,NMAX,KMAX,HGRID,A2,A3,A4,A5,A6,A7
 110    FORMAT(3I3,7E8.1)
       READ120,(NGRID(I),THETA(I),I=1,4)
 120    FORMAT(4(I3,E12.4))
       PRINT10,N1,NMAX,KMAX,HGRID,A2,A3,A4,A5,A6,A7,(NGRID(I),THETA(I),
     1I=1,4)
  10    FORMAT(1H1120X  8HINPUT.../
     1 18X2HN1,1X4HNMAX,1X4HKMAX,9X5HHGRID,12X2HA2,12X2HA3,12X2HA4,
     2 12X2HA5,12X2HA6,12X2HA7/15X3I5,7E14.4//
     3 21X4HN(1),10X8HTHETA(1),6X4HN(2),10X8HTHETA(2),
     4  6X4HN(3),10X8HTHETA(3),6X4HN(4),10X8HTHETA(4)/15X4(I10,E18.6))
        RETURN
          END
C260584              DAUX
      SUBROUTINE DAUX
       COMMON SCRACH,T,NEQ,N1,NMAX,KMAX,HGRID,A2,A3,A4,A5,A6,A7,
     1  NGRID,THETA,W,H,P,A,B,X,U,Y,V
       DIMENSION SCRACH(64),T(250),NGRID(4),THETA(4),W(4,500),H(4,4,500)
     1  ,P(4,500),A(16,16),B(16,1)
C
      XX=X**2
      YY=Y**2
      R=XX+YY
      S=R**3
      D=SQRTF(R**2*S)
      E=SQRTF(S)
      AA=3.*X*Y/D
      BB=(2.*XX-YY)/D
      CC=(2.*YY-XX)/D
        N=23
      DO 1 I=1,4
        N=N+1
      T(N)=T(N-19)
      T(N+1)=T(N-20)*BB + T(N-18)*AA
      T(N+2)=T(N-17)
      T(N+3)=T(N-20)*AA + T(N-18)*CC
    1   N=N+3
      T(40)=T(21)
      T(41)=T(20)*BB + T(22)*AA - 3.*X/E
      T(42)=T(23)
      T(43)=T(20)*AA + T(22)*CC - 3.*Y/E
        RETURN
          END
C260583              START(W)
```

```
      SUBROUTINE START(W)
       COMMON SCRACH,T,NEQ,N1,NMAX,KMAX,HGRID,A2,A3,A4,A5,A6,A7,
     1  NGRID,THETA,W,H,P,A,B,X,U,Y,V
       DIMENSION SCRACH(64),T(250),NGRID(4),THETA(4),W(4,500),H(4,4,500)
     1  ,P(4,500),A(16,16),B(16,1)
            X=1.  X'=0.  Y=0.  Y'=0
      K=0
      PRINT40,K
   40   FORMAT(1H1 65X 9HITERATION,I3//26X4HGRID,14X1HX,19X2HX',18X1HY,
     1  19X2HY')
      DO 2 N=1,NMAX
        W(1,N)=1.
        DO 1 I=2,4
    1   W(I,N)=0.
    2 PRINT50,N,(W(I,N),I=1,4)
   50   FORMAT(20X,I10,4E20.6)

        RETURN
          END
```

Приложение 6

Программа для решения задач кардиологии

```
CMAIN2          HEART MODEL - INVERSE PROBLEM - 2
C
      COMMON NMAX,DELTA,TAU(5),SIGMA(5),ZKAPPA(5),DCOS(5,3),XCHAMB(5,3),
     1  XOBSVR(4,3),ADISTN(3,5),F(5,81),BOBS(3,81),TIME(81),T(1323)
      COMMON KMAX, ASIGMA(5),AKAPPA(5),P(5,81),H(5,10,81),PS(5),
     1  HS(5,10),FPREV(5),AMATRX(243,10),BVECTR(243),EMATRX(50,50),
     2  FVECTR(50),C(10),V(3,81),NCHANG,ICHANG(5),FCHANG(5),IFLAG
     3  ,SCHANG(5)
C
C
C            DISTANCE MODEL
C            5 DIPOLES OR CHAMBERS
C            GAUSSIAN FUNCTION FOR MAGNITUDE OF MOMENTS
C               TAU,SIGMA,ZKAPPA = TIME OF MAX., HALF-WIDTH, MAX. MOMENT
C            DCOS(I,J) = DIREC. COSINE OF DIPOLE I, AXIS J
C            XCHAMB(I,J) = LOCATION OF DIPOLE I, COORDINATE J
C            XOBSVR(I,J) = LOCATION OF OBSERVER I, COORDINATE J
C               J=1 = X-AXIS OR X-DIRECTION
C               J=2 = Y-AXIS OR Y-DIRECTION
C               J=3 = Z-AXIS OR Z-DIRECTION
C
C
C
      CALL FPST
    1 READ1000,NMAX
      PRINT899
      PRINT900,NMAX
      READ1001,DELTA
      PRINT901,DELTA
C
C            INPUT DIPOLE PARAMETERS
C
      DO 2 N=1,5
    2 READ1002,I,TAU(I),SIGMA(I),ZKAPPA(I)
C
      DO 3 N=1,5
    3 READ1003,I,(DCOS(I,J),J=1,3),(XCHAMB(I,J),J=1,3)
      PRINT902
      DO 4 I=1,5
    4 PRINT903,I,TAU(I),SIGMA(I),ZKAPPA(I),(XCHAMB(I,J),J=1,3),
     1          (DCOS(I,J),J=1,3)
C
C            INPUT OBSERVER LOCATIONS
C            COMPUTE DIPOLE-OBSERVER COEFFICIENTS
C
      DO 5 N=1,4
    5 READ1003,I,(XOBSVR(I,J),J=1,3)
      DO 6 I=1,5
        CONUM=0.0
        CODEN=0.0
        DO 61 L=1,3
        YX=XOBSVR(4,L) - XCHAMB(I,L)
        CONUM=CONUM + DCOS(I,L)*YX
   61   CODEN=CODEN + YX**2
```

```
        CODEN=SQRT(CODEN**3)
      CZEROI=CONUM/CODEN
      DO 6 J=1,3
        CJNUM=0.0
        CJDEN=0.0
        DO 62 L=1,3
        YX=XOBSVR(J,L) - XCHAMB(I,L)
        CJNUM=CJNUM + DCOS(I,L)*YX
   62   CJDEN=CJDEN + YX**2
        CJDEN=SQRT(CJDEN**3)
    6 ADISTN(J,I)= CJNUM/CJDEN - CZEROI
C
      PRINT904
      DO 7 J=1,3
    7 PRINT905,J,(XOBSVR(J,L),L=1,3),(ADISTN(J,I),I=1,5)
      J=4
      PRINT905,J,(XOBSVR(J,L),L=1,3)
C
C
C           INPUT CHANGES (SIGMA, KAPPA)
C
   70 PRINT908
      READ1000,NCHANG
      PRINT900,NCHANG
      IF(NCHANG.LE.0)GO TO 80
      READ1000,(ICHANG(I),I=1,NCHANG)
      PRINT900,(ICHANG(I),I=1,NCHANG)
      READ1001,(SCHANG(I),FCHANG(I),I=1,NCHANG)
      PRINT901,(SCHANG(I),FCHANG(I),I=1,NCHANG)
C
C           STORE CHANGES
      DO 71 I=1,NCHANG
      L=ICHANG(I)
      SSIG=SIGMA(L)
      SIGMA(L)=SCHANG(I)
      SCHANG(I)=SSIG
      SKAP=ZKAPPA(L)
      ZKAPPA(L)=FCHANG(I)
   71 FCHANG(I)=SKAP
C
C
C
C
C           INITIAL INTEGRATION STEP
C
   80 IFLAG=1
        K=1
        L=3
      DO 8 I=1,5
        L=L+1
    8 T(L)=ZKAPPA(I)*EXP(-0.5*(TAU(I)/SIGMA(I))**2)
      DO 81 I=1,5
        L=L+1
   81 T(L)=SIGMA(I)
```

```
      T(2)=0.0
      T(3)=DELTA
      CALL INTS(T,10,2,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0)
        L=3
      DO 82 I=1,5
        L=L+1
   82 F(I,K)=T(L)
C           OBSERVATION
      DO 83 J=1,3
      BOBS(J,K)=0.0
      DO 83 I=1,5
   83 BOBS(J,K)=BOBS(J,K) + ADISTN(J,I)*F(I,K)
      TIME(K)=T(2)
      PRINT906
      PRINT907,TIME(K),(F(I,K),I=1,5),(BOBS(J,K),J=1,3)
C
C
C           INTEGRATE
C
      DO 100 K=2,NMAX
      CALL INTM
        L=3
      DO 91 I=1,5
        L=L+1
   91 F(I,K)=T(L)
C
      DO 92 J=1,3
      BOBS(J,K)=0.0
      DO 92 I=1,5
   92 BOBS(J,K)=BOBS(J,K) + ADISTN(J,I)*F(I,K)
      TIME(K)=T(2)
      PRINT907,TIME(K),(F(I,K),I=1,5),(BOBS(J,K),J=1,3)
  100 CONTINUE
C
C
C           PHASE II - DETERMINATION OF HEART PARAMETERS
C
      CALL PHASE2
C
C
C
      IF(NCHANG.LE.0) GO TO 102
C
C           RESTORE ZKAPPA
      DO 101 I=1,NCHANG
      L=ICHANG(I)
      SIGMA(L)=SCHANG(I)
  101 ZKAPPA(L)=FCHANG(I)
C
C           NEXT PROBLEM
  102 READ1000,IGO
      PRINT900,IGO
      GO TO(1,70),IGO
C
```

```
C
C            FORMATS
C
 1000 FORMAT(6I12)
 1001 FORMAT(6E12.6)
  899 FORMAT(1H1 4X38HHEART MODEL - INVERSE PROBLEM -PHASE I //)
  900 FORMAT(6I20)
  901 FORMAT(6E20.6)
 1002 FORMAT(I10,3F10.4)
 1003 FORMAT(I10,6F10.4)
  902 FORMAT(7H0DIPOLE,3X,7X3HTAU,5X5HSIGMA,5X5HKAPPA,5X,3X7HX-COORD,
     1        3X7HY-COORD,3X7HZ-COORD,5X,4X 6HX-DCOS,4X6HY-DCOS,
     2        4X6HZ-DCOS //)
  903 FORMAT(I5,5X3F10.4,5X3F10.4,5X3F10.4)
  904 FORMAT(10H0OBSERVER ,5X, 3X7HX-COORD,3X7HY-COORD,3X7HZ-COORD,5X,
     1        11X1H1,11X1H2,11X1H3,11X1H4,11X1H5//)
  905 FORMAT(I10,5X3F10.4, 5X5F12.6)
  906 FORMAT(1H08X1HT,5X, 5X5HF1(T),5X5HF2(T),5X5HF3(T),5X5HF4(T),
     1        5X5HF5(T),5X,7X5HV1(T),7X5HV2(T),7X5HV3(T)//)
  907 FORMAT(F10.4,5X5F10.4,5X3F12.6)
  908 FORMAT(1H1 4X13HABNORMALITIES //)
      END
$IBFTC PHASE2
      SUBROUTINE PHASE2
      COMMON NMAX,DELTA,TAU(5),SIGMA(5),ZKAPPA(5),DCOS(5,3),XCHAMB(5,3),
     1  XOBSVR(4,3),ADISTN(3,5),F(5,81),BOBS(3,81),TIME(81),T(1323)
      COMMON KMAX, ASIGMA(5),AKAPPA(5),P(5,81),H(5,10,81),PS(5),
     1  HS(5,10),FPREV(5),AMATRX(243,10),BVECTR(243),EMATRX(50,50),
     2  FVECTR(50),C(10),V(3,81),NCHANG,ICHANG(5),FCHANG(5),IFLAG
C
C            SOLVE THE INVERSE PROBLEM
C
C            ASIGMA = APPROX. TO SIGMA
C            AKAPPA = APPROX. TO ZKAPPA
C            P(I,K) = PARTICULAR SOLU. DIPOLE I, TIME K (MOMENT)
C            H(I,J,K)=HOMOGENEOUS SOLU. J,...
C            PS(I)  = PARTICULAR SOLU. DIPOLE I (SIGMA COMPONENT)
C            HS(I,J)= HOMOGENEOUS SOLU. J,...
C            V(J,K) = POTENTIAL DIFFERENCE OBSERVED BY ELECTRODE J
C                     AT TIME K, AS COMPUTED IN PHASE II. COMPARE
C                     WITH BOBS(J,K) OF PHASE I.
C
C
C            INPUT
      PRINT899
      READ1000,KMAX
      PRINT900,KMAX
      PRINT904
      DO 1 N=1,5
      READ1002,I,ASIGMA(I),AKAPPA(I)
    1 PRINT905,I,ASIGMA(I),AKAPPA(I)
C            INITIAL APPROXIMATION
        K1=0
      PRINT909,K1
```

```
      CALL APPRX1
C
C          *HIGHER APPROXIMATIONS
C
      IFLAG=2
      DO 200 K1=1,KMAX
        K=1
      DO 2 I=1,5
      P(I,1)=0.0
      PS(I) =0.0
      DO 3 J=1,10
      H(I,J,1)=0.0
    3 HS(I,J)=0.0
      H(I,I,1)=1.0
        I5=I+5
    2 HS(I,I5)=1.0
        L=3
      DO 4 I=1,5
        L=L+1
    4 T(L)=P(I,1)
      DO 5 I=1,5
        L=L+1
    5 T(L)=PS(I)
      DO 6 J=1,10
      DO 7 I=1,5
        L=L+1
    7 T(L)=H(I,J,1)
      DO 6 I=1,5
        L=L+1
    6 T(L)=HS(I,J)
      T(2)=0.0
      T(3)=DELTA
      DO 8 I=1,5
    8 FPREV(I)=F(I,K)
C
      CALL INTS(T,110,2,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0)
C
      DO 9 K=2,NMAX
      CALL INTM
        L=3
      DO 10 I=1,5
        L=L+1
   10 P(I,K)=T(L)
        L=L+5
      DO 11 J=1,10
      DO 12 I=1,5
        L=L+1
   12 H(I,J,K)=T(L)
   11   L=L+5
      DO 9 I=1,5
    9 FPREV(I)=F(I,K)
C          COMPUTE NEW APPROXIMATION
C
      PRINT909,K1
```

```
  200  CALL APPRX2
C
      RETURN
C
  899 FORMAT(1H1 4X 8HPHASE II //)
 1000 FORMAT(6I12)
  900 FORMAT(6I20)
 1002 FORMAT(I10,10X,5F10.4)
  905 FORMAT(I10,2F10.4)
  904 FORMAT(1H0,8X1HI,5X5HSIGMA,5X5HKAPPA//)
  909 FORMAT(1H156X13HAPPROXIMATION,I3//)
      END
$IBFTC APPRX1
      SUBROUTINE APPRX1
      COMMON NMAX,DELTA,TAU(5),SIGMA(5),ZKAPPA(5),DCOS(5,3),XCHAMB(5,3),
     1  XOBSVR(4,3),ADISTN(3,5),F(5,81),BOBS(3,81),TIME(81),T(1323)
      COMMON KMAX, ASIGMA(5),AKAPPA(5),P(5,81),H(5,10,81),PS(5),
     1  HS(5,10),FPREV(5),AMATRX(243,10),BVECTR(243),EMATRX(50,50),
     2  FVECTR(50),C(10),V(3,81),NCHANG,ICHANG(5),FCHANG(5),IFLAG
C
C           PRODUCE INITIAL APPROXIMATION
C           INTEGRATE NONLINEAR D.E.
      IFLAG=1
        K=1
        L=3
      DO 8 I=1,5
        L=L+1
    8 T(L)=AKAPPA(I)*EXP(-0.5*(TAU(I)/ASIGMA(I))**2)
      DO 81 I=1,5
        L=L+1
   81 T(L)=ASIGMA(I)
      T(2)=0.0
      T(3)=DELTA
      CALL INTS(T,10,2,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0)
        L=3
      DO 82 I=1,5
        L=L+1
   82 F(I,K)=T(L)
C           POTENTIAL COMPUTED
      DO 83 J=1,3
      V(J,K)=0.0
      DO 83 I=1,5
   83 V(J,K)=V(J,K) + ADISTN(J,I)*F(I,K)
      PRINT906
      PRINT907,TIME(K),(F(I,K),I=1,5),(V(J,K),J=1,3)
C
C
C           INTEGRATE
      DO 100 K=2,NMAX
      CALL INTM
        L=3
      DO 91 I=1,5
        L=L+1
   91 F(I,K)=T(L)
```

```
C
      DO 92 J=1,3
      V(J,K)=0.0
      DO 92 I=1,5
   92 V(J,K)=V(J,K) + ADISTN(J,I)*F(I,K)
      PRINT907,TIME(K),(F(I,K),I=1,5),(V(J,K),J=1,3)
  100 CONTINUE
      RETURN
C
  906 FORMAT(1H0BX1HT,5X,5X5HF1(T),5X5HF2(T),5X5HF3(T),5X5HF4(T),
     1        5X5HF5(T),5X,7X5HV1(T),7X5HV2(T),7X5HV3(T)///)
  907 FORMAT(F10.4,5X5F10.4,5X3F12.6)
      END
$IBFTC APPRX2
      SUBROUTINE APPRX2
      DIMENSION PIVOT(50),INDEX(50,2),IPIVOT(50),EM(50,50),FV(50,1)
      COMMON NMAX,DELTA,TAU(5),SIGMA(5),ZKAPPA(5),DCOS(5,3),XCHAMB(5,3),
     1  XOBSVR(4,3),ADISTN(3,5),F(5,81),BOBS(3,81),TIME(81),T(1323)
      COMMON KMAX, ASIGMA(5),AKAPPA(5),P(5,81),H(5,10,81),PS(5),
     1  HS(5,10),FPREV(5),AMATRX(243,10),BVECTR(243),EMATRX(50,50),
     2  FVECTR(50),C(10),V(3,81),NCHANG,ICHANG(5),FCHANG(5),IFLAG
C
C           APPLY BOUNDARY CONDITIONS
C           SOLVE LINEAR EQUATIONS
C           COMPUTE NEW APPROXIMATION
C
C           SET UP MATRIX, VECTOR
        I=0
      DO 3 J=1,3
      DO 3 K=1,81
        I=I+1
      DO 1 L=1,10
      AMATRX(I,L)=0.0
      DO 1 M=1,5
    1 AMATRX(I,L)=AMATRX(I,L) + ADISTN(J,M)*H(M,L,K)
      BVECTR(I)=BOBS(J,K)
      DO 2 M=1,5
    2 BVECTR(I)=BVECTR(I) - ADISTN(J,M)*P(M,K)
    3 CONTINUE
C
      DO 6 I=1,10
      DO 4 J=1,10
      EMATRX(I,J)=0.0
      DO 4 L=1,243
    4 EMATRX(I,J)=EMATRX(I,J) + AMATRX(L,I)*AMATRX(L,J)
      FVECTR(I)=0.0
      DO 5 L=1,243
    5 FVECTR(I)=FVECTR(I) + AMATRX(L,I)*BVECTR(L)
    6 CONTINUE
C
C           INVERT MATRIX
      DO 61 I=1,50
      FV(I,1)=FVECTR(I)
      DO 61 J=1,50
```

```
   61  EM(I,J)=EMATRX(I,J)
       CALL MATINV(EM,10,FV,1,DETERM,PIVOT,INDEX,IPIVOT)
       DO 62 I=1,50
   62  FVECTR(I)=FV(I,1)
C
C            NEW APPROXIMATION
       PRINT904
       DO 7 I=1,10
    7  C(I)=FVECTR(I)
       DO 8 I=1,5
       F(I,1)=C(I)
       ASIGMA(I)=C(I+5)
       AKAPPA(I)=F(I,1)*EXP(0.5*(TAU(I)/ASIGMA(I))**2)
    8  PRINT905,I,ASIGMA(I),AKAPPA(I)
       PRINT906
         K=1
       DO 83 J=1,3
       V(J,K)=0.0
       DO 83 I=1,5
   83  V(J,K)=V(J,K) + ADISTN(J,I)*F(I,K)
       PRINT907,TIME(K),(F(I,K),I=1,5),(V(J,K),J=1,3)
C
       DO 100 K=2,NMAX
       DO 9 I=1,5
       F(I,K)=P(I,K)
       DO 9 J=1,10
    9  F(I,K)=F(I,K) + C(J)*H(I,J,K)
       DO 93 J=1,3
       V(J,K)=0.0
       DO 93 I=1,5
   93  V(J,K)=V(J,K) + ADISTN(J,I)*F(I,K)
  100  PRINT907,TIME(K),(F(I,K),I=1,5),(V(J,K),J=1,3)
       RETURN
C
  904  FORMAT( 1H0,8X1HI, 5X 5HSIGMA, 5X 5HKAPPA//)
  905  FORMAT(I10,2F10.4)
  906  FORMAT(1H08X1HT,5X, 5X5HF1(T),5X5HF2(T),5X5HF3(T),5X5HF4(T),
      1        5X5HF5(T),5X,7X5HV1(T),7X5HV2(T),7X5HV3(T)//)
  907  FORMAT(F10.4,5X5F10.4,5X3F12.6)
       END
$IBFTC DAUX
       SUBROUTINE DAUX
       COMMON NMAX,DELTA,TAU(5),SIGMA(5),ZKAPPA(5),DCOS(5,3),XCHAMB(5,3),
      1  XOBSVR(4,3),ADISTN(3,5),F(5,81),BOBS(3,81),TIME(81),T(1323)
       COMMON KMAX, ASIGMA(5),AKAPPA(5),P(5,81),H(5,10,81),PS(5),
      1  HS(5,10),FPREV(5),AMATRX(243,10),BVECTR(243),EMATRX(50,50),
      2  FVECTR(50),C(10),V(3,81),NCHANG,ICHANG(5),FCHANG(5),IFLAG
       DIMENSION XF(5),XSG(5)
C
C
       GO TO (1,2),IFLAG
C
C
C            NONLINEAR EQUATIONS
```

```
C
    1     L=3
          M=13
        DO 10 I=1,5
          L=L+1
          M=M+1
   10   T(M)=-(T(2)-TAU(I))*T(L)/(T(L+5)**2)
        DO 11 I=1,5
          M=M+1
   11   T(M)=0.0
        RETURN
C
C             LINEAR EQS. FOR P,  H
CPARTICULAR
    2     L=3
        DO 3 I=1,5
          L=L+1
    3   XF(I)=T(L)
        DO 4 I=1,5
          L=L+1
    4   XSG(I)=T(L)
C
          M=113
        DO 5 I=1,5
          M=M+1
        TX=-(T(2)-TAU(I))/(ASIGMA(I)**2)
    5   T(M)=TX*(XF(I) + XSG(I)*(-2.0*FPREV(I)/ASIGMA(I)) + 2.0*FPREV(I))
        DO 6 I=1,5
          M=M+1
    6   T(M)=0.0
C
C
CHOMOGENEOUS
        DO 13 J=1,10
C
        DO 7 I=1,5
          L=L+1
    7   XF(I)=T(L)
        DO 8 I=1,5
          L=L+1
    8   XSG(I)=T(L)
C
        DO 9 I=1,5
          M=M+1
        TX=-(T(2)-TAU(I))/(ASIGMA(I)**2)
    9   T(M)=TX*(XF(I) + XSG(I)*(-2.0*FPREV(I)/ASIGMA(I)) )
        DO 12 I=1,5
          M=M+1
   12   T(M)=0.0
C
   13   CONTINUE
        RETURN
        END
```

ОГЛАВЛЕНИЕ

Редактор *Л. Б. Штейнпресс*
Художник *Л. А. Грибов*
Художественный редактор *В. И. Шаповалов*
Технический редактор *Т. В. Чечик*
Корректор *О. К. Румянцева*

Сдано в производство 20/II 1968 г.
Подписано к печати 1/VII 1968 г.
Бумага глуб. печ. 60×90 $^{1}/_{16}$= 5,75 бум. л.
11,5 печ. л. Уч.-изд. л. 9,48. Изд. № 1/4426.
Цена 86 к. Зак. 1113.

ИЗДАТЕЛЬСТВО „МИР"
Москва, 1-й Рижский пер., 2

Ленинградская типография № 2
имени Евгении Соколовой
Главполиграфпрома
Комитета по печати
при Совете Министров СССР.
Измайловский проспект, 29.