SOLUTION OF EQUATIONS
and
SYSTEMS OF EQUATIONS

A. M. OSTROWSKI

UNIVERSITY OF BASEL, BASEL, SWITZERLAND
ACADEMIC PRESS · NEW YORK AND LONDON · 1960

А. М. ОСТРОВСКИЙ

РЕШЕНИЕ УРАВНЕНИЙ И СИСТЕМ УРАВНЕНИЙ

Перевод с английского
Л. З. РУМШИСКОГО и Б. Л. РУМШИСКОГО

ИЗДАТЕЛЬСТВО ИНОСТРАННОЙ ЛИТЕРАТУРЫ
Москва 1963

АННОТАЦИЯ

Эта небольшая книжка знакомит читателя с математическим обоснованием и исследованием методов численного анализа. Наряду с разработкой новых эффективных методов приближенного решения уравнений автор проводит глубокие и тонкие исследования сходимости уже известных методов (метода Ньютона, метода Стеффенсена и др.). Большой интерес представляет предлагаемый автором новый способ сравнения численных методов решения уравнений, основанный на введенном им «индексе эффективности».

Исследования А. М. Островского касаются и такого важного для практики вопроса, как округление при вычислениях; здесь автором также получены интересные результаты.

В основном книга рассчитана на математиков-вычислителей. Представителям же других отраслей математики она, по-видимому, будет интересна как образец проникновения «чистой» математики в область «прикладной» математики. Книга окажется безусловно полезной также физикам и инженерам, применяющим численные методы.

Редакция литературы по математическим наукам

ПРЕДИСЛОВИЕ ПЕРЕВОДЧИКОВ

За последние годы в Советском Союзе издано много книг, посвященных различным вопросам численного анализа. Однако авторы этих книг обычно мало занимаются математическим исследованием численных методов. Поэтому предлагаемая читателю книга А. М. Островского представляет несомненный интерес. Островский, наряду с разработкой новых эффективных методов приближенного решения уравнений, проводит глубокие и тонкие исследования сходимости уже известных методов (метод Ньютона, regula falsi, метод Стеффенсена и др.). При этом он получает ряд новых, а иногда даже неожиданных результатов. Например, Островский показал, что так называемый метод секущих в некоторых отношениях более эффективен, чем метод Ньютона (гл. 3,5). Другой пример: в вычислительной практике для получения двусторонних приближений к корню уравнения часто применяется комбинация методов Ньютона и regula falsi («метод хорд и касательных»). Среди вычислителей распространено мнение, что при этом приближения, получаемые по методу Ньютона, имеют более высокий порядок точности, чем приближения по requla falsi. Островскому удалось показать, что скорость сходимости с обеих сторон одинакова (теорема 10.1).

Особенностью исследования автора является систематическое применение обратной интерполяции как для получения новых методов решения уравнений, так и особенно для получения более точных оценок погрешности рассматриваемых методов. Заслуживает серьезного внимания предлагаемый автором новый способ сравнения численных методов решения уравнений, основанный на введенном им «индексе эффективности».

Весьма важны для практики разработанные автором способы улучшения сходимости численных методов, в частности способы улучшения методов со сверхлинейной сходимостью. Исследования Островского охватывают и вопросы округления при вычислениях; здесь автором получены очень глубокие результаты.

Значительная часть книги посвящена установлению критериев сходимости различных итерационных процессов, исследованию случая кратных корней и другим вопросам общей теории итерации. «Попутно» автор получил ряд интересных результатов из области «чистой математики».

Несмотря на такое обилие и разнообразие материала, книга написана очень компактно. Почти все исследования автора проведены сравнительно элементарными методами классического анализа. Поэтому основная часть книги доступна для читателя, знакомого с высшей математикой лишь в объеме обычной программы втуза. Для специалистов-математиков книга будет интересна как образец проникновения методов «чистой» математики в область «прикладной», вычислительной математики.

Книга Островского написана очень неровно. Отдельные доказательства растянуты; в приложения включен весьма неравноценный материал.

При переводе мы старались сохранить стиль автора и максимально точно передать его суждения о ценности тех или иных методов. Однако, чтобы сделать книгу более доступной, мы упростили выкладки при доказательстве ряда теорем и опустили некоторые малозначительные детали. Кроме того, исправлены имевшиеся в оригинале отдельные неточности и опечатки.

Л. З. Румшиский,
Б. Л. Румшиский

ПРЕДИСЛОВИЕ АВТОРА

Эта книга возникла из курса лекций, прочитанного по приглашению Национального бюро стандартов (НБС) летом 1952 года в Вашингтонском университете. Конспект лекций, записанный М. Тиксоном, был под номером CL52-2 размножен в небольшом числе экземпляров для служебного пользования в НБС.

Так как существенную часть этого курса составляли неопубликованные результаты, мне представилось целесообразным опубликовать весь курс отдельной книгой. С этой целью весь курс был заново переработан и полностью переписан; к лекциям было добавлено большое количество нового материала. Кроме того, я написал 11 приложений, посвященных проблемам, которые не могут быть рассмотрены в основной части книги. Однако изложение все еще далеко от полноты.

В то время как лекции были рассчитаны на более или менее неподготовленного слушателя, книга в настоящем своем виде использует методы и приемы доказательства, более соответствующие уровню знаний лиц, окончивших высшее учебное заведение. На мой взгляд, практическая важность численных методов возросла до такой степени, что математикам стоит использовать все свое искусство, чтобы дать полное обоснование этим методам. Я надеюсь, что эта книга до известной степени поможет перекинуть мост над пропастью, разделяющей «чистую» и «прикладную» математику.

Заканчивая эту книгу, я пользуюсь возможностью выразить благодарность Джону Х. Кертиссу, бывшему в то время руководителем математического отдела НБС, по инициативе которого были прочитаны эти лекции. Несомненно, заслугой Кертисса является создание той исключительно благоприятной атмосферы, в которой исследовательские

группы НБС могли развивать свою активность. Беседы с Ольгой Таусской-Тодд и Джоном Тоддом всегда были исключительно вдохновляющими. Я выражаю также благодарность М. Тиксону, впервые законспектировавшему эти лекции, и В. Ф. Кэйхиллу, который был в то время моим ассистентом. При подготовке этих лекций большую помощь мне оказали вычислительные центры НБС (находящиеся в Вашингтоне и Лос-Анжелосе). Эта помощь была особенно полезна для проверки на практике различных методов. Наконец, я должен поблагодарить моих ассистентов в Базеле Б. Марцетта, С. Кристеллера, Т. Витцеманна и Р. Бюрки, а также Е. В. Хэйнсуорта и В. Гочи из НБС за большую помощь, которую они оказали мне при подготовке рукописи этой книги, и Пьера Бандере и Говарда Белла за помощь в исправлении корректур.

А. М. Островский

Июнь 1960 г.

ГЛАВА I

ВВЕДЕНИЕ. ОСТАТОЧНЫЕ ЧЛЕНЫ ИНТЕРПОЛЯЦИОННЫХ ФОРМУЛ

1. В этой книге мы рассмотрим методы численного решения уравнений: сначала одного уравнения с одним неизвестным, затем системы уравнений с соответствующим числом неизвестных.

Бо́льшая часть рассматриваемых методов основана на интерполяции. Поэтому мы начнем с изучения остаточных членов интерполяционных формул.

В дальнейшем J_x будет обозначать сегмент оси x, а (J_x) — *внутренность* этого сегмента, т. е. *открытый интервал*.

Лемма о функциях с несколькими корнями

2. Лемма 1.1. Если определенная на J_x функция $F(x)$ обращается в нуль в $n+1$-й точке $x_\nu \in J_x$ $(\nu = 1, \ldots, n+1)$ и если $F^{(n)}(x)$ существует на J_x, то найдется точка $\xi \in J_x$, в которой $F^{(n)}(\xi) = 0$. Точка ξ лежит даже в (J_x), если только не все x_ν совпадают с одним концом J_x.

Замечание. Корни x_ν функции $F(x)$ могут быть *кратными*. Каждый кратный корень надо считать столько раз, какова его кратность.

Лемма 1.1 является обобщением теоремы Ролля. Как пример исключительного случая, отмеченного в последнем предложении леммы, укажем функцию

$$F(x) = x^{n+1} \text{ на сегменте } J_x : 0 \leqslant x \leqslant 1.$$

Здесь $\xi = 0$, так что ξ не принадлежит $(J_x) : 0 < x < 1$.

3. Доказательство. Покажем сначала, что функция $g(x) \equiv F'(x)$ имеет на J_x не менее n корней.

По теореме Ролля *между* каждыми двумя *различными* корнями функции $F(x)$ лежит по крайней мере один корень производной $F'(x)$. Поэтому если все x_ν $(\nu = 1, \ldots, n+1)$ — простые корни функции $F(x)$, то $g(x)$ имеет не менее n различных корней, разделяющих корни x_ν.

В общем случае обозначим через k число *различных* кратных корней функции $F(x)$ на J_x (причем простой корень считаем корнем кратности 1). Каждый корень кратности μ функции $F(x)$ является корнем кратности $\mu - 1$ функции $g(x)$, т. е. каждый кратный корень функции $F(x)$ теряет единицу своей кратности в $g(x)$. Поэтому $g(x)$ имеет $(n+1) - k$ корней, совпадающих с кратными корнями $F(x)$. Кроме того, $g(x)$ имеет не менее чем $(k-1)$ «новых» корней, разделяющих, по теореме Ролля, k различных корней $F(x)$. Итак, $g(x)$ имеет не менее чем $(n+1) - k + (k-1) = n$ корней на J_x.

4. Рассмотрим теперь два случая.

a) $x_1 = x_2 = \ldots = x_{n+1}$. В этом случае лемма верна, так как x_1 есть корень $F^{(n)}(x)$, т. е. $\xi = x_1$.

b) Среди корней x_ν есть хотя бы два различных. В этом случае мы докажем лемму по индукции.

При $n = 1$ лемма совпадает с теоремой Ролля и потому верна. Предположим теперь, что лемма верна для всех чисел, меньших, чем $n+1$. Тогда лемма верна для $g(x)$, так как $g(x)$ имеет по крайней мере n корней на J_x; следовательно, существует точка $\xi \in J_x$, в которой $g^{(n-1)}(\xi) = F^{(n)}(\xi) = 0$. Так как при этом хотя бы один из корней функции $g(x)$ лежит в (J_x), то точка ξ лежит в (J_x). Лемма доказана.

Теорема об отношении функций, имеющих общие корни

5. Теорема 1.1. Пусть $f(x)$, $g(x)$ определены и n раз дифференцируемы на J_x, причем $g^{(n)}(x)$ не обращается в нуль на (J_x). Пусть $f(x)$ и $g(x)$ имеют n общих корней $x_1, x_2, \ldots, x_n \in J_x$, причем некоторое число повторяется в этом перечне k раз, если оно является не менее чем k-кратным корнем как $f(x)$, так и $g(x)$. Тогда для

каждого $x_0 \neq x_\nu$ из J_x существует такая точка $\xi \in (J_x)$, что

$$\frac{f(x_0)}{g(x_0)} = \frac{f^{(n)}(\xi)}{g^{(n)}(\xi)}, \quad \xi \in (J_x), \quad (x_0 \neq x_\nu;\ \nu = 1, \ldots, n). \tag{1.1}$$

6. Доказательство. Прежде всего, x_0 не является корнем $g(x)$; в противном случае функция $g(x)$ имела бы $n+1$ корень в J_x и по лемме 1.1 в некоторой точке $\xi \in (J_x)$ было бы $g^{(n)}(\xi) = 0$ вопреки условию теоремы.

Положим $\lambda = \frac{f(x_0)}{g(x_0)}$ и рассмотрим $F(x) = f(x) - \lambda g(x)$. Функция $F(x)$ удовлетворяет условиям леммы 1.1: она n раз дифференцируема на J_x вместе с $f(x)$ и $g(x)$; ее корнями являются все точки x_ν $(\nu = 1, \ldots, n)$ и x_0, причем по крайней мере два из этих корней различны (по условию $x_0 \neq x_1$). Поэтому существует точка $\xi \in (J_x)$, в которой

$$F^{(n)}(\xi) = f^{(n)}(\xi) - \lambda g^{(n)}(\xi) = 0,$$

т. е.

$$\frac{f^{(n)}(\xi)}{g^{(n)}(\xi)} = \lambda = \frac{f(x_0)}{g(x_0)}, \text{ ч. т. д.}$$

В доказанной теореме и в лемме 1.1 предположения о том, что функции $F(x)$, $f(x)$, $g(x)$ имеют n производных только в (J_x) и непрерывны на J_x, было бы достаточно во всех случаях, за исключением того, когда один из концов сегмента J_x является кратным корнем. В последнем случае рассматриваемые функции должны иметь на этом конце сегмента столько производных, сколько их требуется в силу кратности соответствующего корня.

Теорема 1.1 позволяет исследовать *остаточные члены* интерполяционных формул.

Интерполирующие функции

7. Пусть $f(x)$ определена в n узлах интерполяции $x_\nu \in J_x$ $(\nu = 1, \ldots, n)$. Определенная на J_x функция $T(x)$ называется *интерполирующей функцией* для $f(x)$ с интерполяционными узлами x_ν, если

$$T(x_\nu) = f(x_\nu) \quad (\nu = 1, \ldots, n). \tag{1.2}$$

Функция $T(x)$ может быть выбрана из различных классов функций, например, из многочленов, тригонометрических функций, рациональных функций от x и т. д.

Среди значений x_ν могут встречаться *равные*, например, запись $x_\nu = a$ $(\nu = 1, \ldots, m)$ означает, что в точке a функция $f(x) - T(x)$ имеет корень кратности m, а соответствующие уравнения (1.2) заменены на

$$T^{(\mu)}(a) = f^{(\mu)}(a) \quad (\mu = 0, 1, \ldots, m-1). \tag{1.3}$$

Так, если $x_1 = x_2 = x_3 = a$, а остальные $n-3$ узла интерполяции попарно различны, то система (1.2) заменяется на

$$\begin{aligned} &T^{(\mu)}(a) = f^{(\mu)}(a) \quad (\mu = 0, 1. 2);\\ &T(x_\nu) = f(x_\nu) \qquad (\nu = 4, 5, \ldots, n). \end{aligned}$$

Если все n узлов интерполяции совпадают, а в качестве $T(x)$ выбран многочлен степени $(n-1)$, то $T(x)$ есть известный многочлен Тейлора.

Общий вид остаточного члена интерполяционной формулы

8. Пусть $f(x)$ определена и n раз дифференцируема на J_x. Пусть $T(x)$ — интерполирующая функция для $f(x)$ с узлами интерполяции $x_1, \ldots, x_n$ из J_x, причем $T(x)$ также n раз дифференцируема на J_x. Пусть, наконец, $g(x)$ — произвольная функция, удовлетворяющая условиям

$$g(x\ ) = 0 \quad (\nu = 1, \ldots, n)$$

и имеющая на J_x производную n-го порядка $g^{(n)}(x)$, не обращающуюся в нуль на J_x.

При этом функции $f(x) - T(x)$ и $g(x)$ удовлетворяют условиям теоремы 1.1, и, следовательно, для каждого $x \neq x_\nu$ $(\nu = 1, \ldots, n)$ существует такая точка ξ (зависящая от x), что

$$\frac{f(x) - T(x)}{g(x)} = \frac{f^{(n)}(\xi) - T^{(n)}(\xi)}{g^{(n)}(\xi)},$$

т. е.

$$f(x) - T(x) = \frac{f^{(n)}(\xi) - T^{(n)}(\xi)}{g^{(n)}(\xi)}\, g(x), \quad \xi \in (J_x). \tag{1.4}$$

В частности, при $g(x) = \prod_{\nu=1}^{n} (x - x_\nu)$ получаем

$$f(x) - T(x) = \frac{f^{(n)}(\xi) - T^{(n)}(\xi)}{n!} \prod_{\nu=1}^{n} (x - x_\nu), \quad \xi \in (J_x). \quad (1.4a)$$

Если $T(x)$ — *интерполяционный многочлен* степени $n-1$, то $T^{(n)}(x) \equiv 0$ и

$$f(x) - T(x) = \frac{f^{(n)}(\xi)}{n!} \prod_{\nu=1}^{n} (x - x_\nu), \quad \xi \in (J_x), \quad (1.5)$$

В случае, когда $x_1 = x_2 = \ldots = x_n$, формула (1.5) дает остаточный член формулы Тейлора.

В случае, когда все узлы интерполяции различны, формула (1.5) дает остаточный член n-точечной интерполяционной формулы Лагранжа.

Интерполяционный многочлен Эрмита

9. Пусть $H(x)$ — интерполяционный многочлен степени $n-1$, и пусть среди заданных n узлов интерполяции имеется k различных: $x_1, x_2, \ldots, x_k$, причем x_1 встречается m_1 раз, x_2 встречается m_2 раз, ..., x_k встречается m_k раз, где $m_1 + m_2 + \ldots + m_k = n$. Полагая

$$f^{(\mu)}(x_i) = H^{(\mu)}(x_i) = a_\mu^{(i)} \quad (\mu = 0, 1, \ldots, m_i - 1; \quad i = 1, \ldots, k),$$

получим следующую таблицу значений:

$$\begin{array}{l}
x_1 : a_0^{(1)}, \quad a_1^{(1)}, \ldots, \quad a_{m_1-1}^{(1)}; \\
x_2 : a_0^{(2)}, \quad a_1^{(2)}, \ldots, \quad a_{m_2-1}^{(2)}; \\
\cdots\cdots\cdots\cdots \\
x_k : a_0^{(k)}, \quad a_1^{(k)}, \ldots, \quad a_{m_k-1}^{(k)}.
\end{array}$$

Этими условиями интерполяционный многочлен $H(x)$ (называемый *интерполяционным многочленом Эрмита*) определяется однозначно. Действительно, если существует другой многочлен $H_1(x)$, удовлетворяющий тем же условиям, то многочлен $H - H_1$ степени не выше $n-1$

должен делиться на многочлен

$$(x-x_1)^{m_1}(x-x_2)^{m_2}\ldots(x-x_k)^{m_k}$$

степени n. Поэтому разность $H-H_1$ должна быть тождественно равна нулю, что и доказывает единственность многочлена $H(x)$.

Остаточный член эрмитовской интерполяции получается из формулы (1.5) в виде

$$f(x)-H(x)=\frac{f^{(n)}(\xi)}{n!}(x-x_1)^{m_1}(x-x_2)^{m_2}\ldots(x-x_k)^{m_k}. \quad (1.6)$$

10. В случае, когда все узлы интерполяции различны, т. е. все $m_i=1$ $(i=1,\ldots,n)$, интерполяционный многочлен Эрмита совпадает с интерполяционным многочленом Лагранжа. Интерполяционный многочлен Лагранжа можно записать следующим образом:

$$L(x)=\sum_{\nu=1}^{n} f(x_\nu)\frac{F(x)}{(x-x_\nu)F'(x_\nu)}, \quad (1.7)$$

где

$$F(x)=\prod_{\nu=1}^{n}(x-x_\nu).$$

ГЛАВА 2

ОБРАТНАЯ ИНТЕРПОЛЯЦИЯ. ПРОИЗВОДНЫЕ ОБРАТНОЙ ФУНКЦИИ. ОДИН УЗЕЛ ИНТЕРПОЛЯЦИИ

Понятие об обратной интерполяции

1. Если число a применяется в качестве приближенного значения числа x, то мы будем писать $a \sqsupset x$; такое же обозначение мы будем применять и при аппроксимации функций.

Возможны два различных подхода к задаче нахождения корней функции $f(x)$. При первом подходе функцию $f(x)$ заменяют интерполирующей функцией $T(x)$ и находят корни уравнения $T(x) = 0$. При этом возникает вопрос, насколько хорошо корни этого нового уравнения аппроксимируют корни уравнения $f(x) = 0$. Известно, что даже малое изменение коэффициентов алгебраического уравнения может существенно изменить его корни (см. приложение I). В дальнейшем мы исследуем этот вопрос более подробно и установим условия, при которых корни уравнения $T(x) = 0$ достаточно хорошо аппроксимируют корни уравнения $f(x) = 0$ (приложения I, II и XI).

Второй подход к задаче нахождения корней связан с рассмотрением *обратной функции*.

Пусть функция $y = f(x)$ определена на J_x, и пусть заданы ее значения в n узлах интерполяции:

$$f(x_\nu) = y_\nu \quad (\nu = 1, 2, \ldots, n). \tag{2.1}$$

Пусть $x = \varphi(y)$ есть функция, обратная к функции $y = f(x)$. Тогда задача нахождения приближенного значения корня функции $f(x)$ сводится к задаче вычисления значения $\varphi(0)$ по известным значениям $\varphi(y_\nu) = x_\nu$ $(\nu = 1, 2, \ldots, n)$.

Построив для $\varphi(y)$ интерполяционный многочлен $T(y)$ так, чтобы было $T(y_\nu) = x_\nu$ $(\nu = 1, \ldots, n)$, вычислим $T(0) \sqsupset \varphi(0)$. При этом оценка погрешности может быть получена из (1.5):

$$\varphi(y) - T(y) = \frac{\varphi^{(n)}(\eta)}{n!} \prod_{\nu=1}^{n} (y - y_\nu);$$

$$\varphi(0) - T(0) = \frac{\varphi^{(n)}(\eta)}{n!} (-1)^n y_1 y_2 \ldots y_n. \tag{2.2}$$

Отсюда видно, что если все узлы интерполяции выбраны близко к корню, то погрешность будет мала.

2. Описанный выше метод известен уже давно, но математики обычно избегают его из-за трудностей исследования обратной функции и ее производных. Однако эти трудности только кажущиеся. Решение уравнения $f(x) = 0$ ищут в определенном интервале, внутри которого, как правило, $f'(x) \neq 0$. При этом условии упомянутые трудности, как мы покажем дальше, устраняются. Если же это условие не выполнено, то не только обратная интерполяция, но и широко распространенные методы не дают результата, что вынуждает обращаться к специальным методам.

Теорема Дарбу о значениях производной

3. Докажем теорему, дающую достаточное обоснование нашим утверждениям.

Теорема 2.1 (Дарбу). Пусть $f(x)$ определена и непрерывна на $J_x: a \leqslant x \leqslant b$. Если производная $f'(x)$ существует на J_x и принимает значения $f'(a) = A$ и $f'(b) = B$, то она принимает на J_x все значения, заключенные между A и B.

Доказательство. Без ограничения общности можно предположить, что $A < B$. Пусть C — некоторое число между A и B; докажем, что $f'(x)$ принимает значение C где-либо в (J_x). Рассмотрим $F(x) = f(x) - Cx$. Из формул

$$F'(x) = f'(x) - C; \quad F'(a) = A - C < 0; \quad F'(b) = B - C > 0$$

видно, что непрерывная функция $F(x)$ имеет отрицательную производную в точке a и положительную в точке b. Следовательно, $F(x)$ принимает в (J_x) значения, меньшие чем $F(a)$ и $F(b)$; поэтому функция $F(x)$ имеет минимум

в некоторой точке ξ, лежащей *внутри* J_x, и ее производная в этой точке обращается в нуль: $f'(\xi) - C = 0$; $f'(\xi) = C$, ч. т. д.

Из теоремы Дарбу следует, что если производная $f'(x)$ отлична от нуля на J_x, то она *сохраняет знак* на J_x (либо $f'(x) > 0$ для всех $x \in J_x$, либо $f'(x) < 0$ для всех $x \in J_x$).

Производные обратной функции

4. Пусть $f(x)$ определена на J_x. Если на J_x существует и отлична от нуля производная $f'(x)$, то $f(x)$ строго монотонна и по известной теореме имеет обратную функцию $x = \varphi(y)$, дифференцируемую на соответствующем y-интервале:

$$\varphi'(y) = \frac{dx}{dy} = \frac{1}{y'} = \frac{1}{f'(x)}. \tag{2.3}$$

Если, кроме того, на J_x существует вторая производная $f''(x) = y''$, то и обратная функция имеет вторую производную, получающуюся дифференцированием формулы (2.3):

$$\varphi''(y) = -\frac{y''}{(y')^3}. \tag{2.4}$$

Получить удобные выражения для высших производных обратной функции значительно сложнее. В дальнейшем, предполагая существование $(n+1)$ производных от $f(x)$, мы получим рекуррентную формулу для соответствующих производных от $\varphi(y)$.

5. Покажем, что

$$\varphi^{(k)}(y) = \frac{X_k}{(y')^{2k-1}} \qquad (k = 1, 2, \ldots, n+1), \tag{2.5}$$

где X_k — *многочлен* относительно $y', y'', \ldots, y^{(k)}$. Это верно для $n = 0, 1$, причем $X_1 = 1$, $X_2 = -y''$. Допустим, что наше утверждение справедливо для первых n производных от $\varphi(y)$; дифференцируя формулу (2.5) при $k = n$ и учитывая, что $\frac{dy'}{dy} = \frac{y''}{y'}$, мы получим

$$\varphi^{(n+1)}(y) = \frac{dX_n}{dy}(y')^{-2n+1} - X_n(2n-1)(y')^{-2n}\frac{y''}{y'}; \tag{2.6}$$

следовательно,

$$X_{n+1} = (y')^{2n+1}\varphi^{(n+1)}(y) = \frac{dX_n}{dx}y' - (2n-1)X_n y''. \qquad (2.7)$$

Наше утверждение доказано. Полученная рекуррентная формула (2.7) может служить для нахождения производных обратной функции.

6. Мы будем писать дальше в этой главе y_ν вместо $y^{(\nu)}$. Покажем, что X_n есть *однородный* многочлен относительно $y_1, y_2, \ldots, y_n$ *размерности* $n-1$, т. е. что

$$X_n = \sum a_{\alpha_1 \alpha_2 \ldots \alpha_n} y_1^{\alpha_1} y_2^{\alpha_2} \ldots y_n^{\alpha_n}, \qquad (2.8)$$

где в каждом члене

$$\alpha_1 + \alpha_2 + \ldots + \alpha_n = n - 1. \qquad (2.9)$$

Это верно для X_1 и X_2. Допустим, что (2.9) верно для некоторого n. Тогда из (2.7) следует, что X_{n+1} будет однородным многочленом размерности n, так как это верно и для второго слагаемого $-(2n-1)X_n \cdot y_2$, и для первого слагаемого

$$\frac{dX_n}{dx}y_1 = \sum_\nu \frac{\partial X_n}{\partial y_\nu} y_{\nu+1} \cdot y_1. \qquad (2.10)$$

Далее, если каждой переменной y_ν приписать *вес* ν, многочлен X_n будет *изобарическим* с *общим весом* $(2n-2)$, т. е. в каждом члене суммы (2.8)

$$\alpha_1 + 2\alpha_2 + 3\alpha_3 + \ldots + n\alpha_n = 2n - 2. \qquad (2.11)$$

Это верно для X_1 и X_2. Допустим, что соотношение (2.11) верно для некоторого n. Тогда общий вес второго) слагаемого в (2.7) на 2 больше, чем общий вес X_n, а процесс (2.10) увеличивает вес каждого члена X_n также на 2.

7. Наконец, в (2.8) есть только один член, содержащий y_n (*старший член*):

$$-y_n y_1^{n-2} \quad (n \geqslant 2), \qquad (2.12)$$

и только один член, не содержащий никаких y_ν с $\nu > 2$ (*младший член*):

$$(-1)^{n-1} 1 \cdot 3 \cdots (2n-3) y_2^{n-1} \quad (n \geqslant 2). \qquad (2.13)$$

Это верно для X_2. Допустим, что наши утверждения верны для некоторого n. Тогда из формул (2.7) и (2.10) видно, что единственный член X_{n+1}, содержащий y_{n+1}, может быть получен только из (2.12) в виде

$$\frac{\partial}{\partial y_n}\left(-y_n y_1^{n-2}\right)\cdot y_{n+1}\cdot y_1 = -y_{n+1}\, y_1^{n-1}.$$

Далее, если младший член в X_n есть (2.13), то производная от него уже содержит y_3; поэтому единственный член в X_{n+1}, зависящий только от y_1 и y_2, получается из второго слагаемого правой части (2.7) умножением (2.13) на $-(2n-1)\,y_2$. Наши утверждения доказаны.

Ниже мы даем таблицу для $X_1, \ldots, X_6$. Явная формула для X_n и исследование частных случаев даны в приложении III.

8. *Таблица* $X_1, \ldots, X_6$:

$$X_1 = 1.$$

$$X_2 = -y_2.$$

$$X_3 = -y_3 y_1 + 3y_2^2.$$

$$X_4 = -y_4 y_1^2 + 10y_3 y_2 y_1 - 15y_2^3.$$

$$X_5 = -y_5 y_1^3 + 15y_4 y_2 y_1^2 + 10y_3^2 y_1^2 - 105y_3 y_2^2 y_1 + 105y_2^4.$$

$$X_6 = -y_6 y_1^4 + 21y_5 y_2 y_1^3 + 35y_4 y_3 y_1^3 - 210y_4 y_2^2 y_1^2 - \\ - 280y_3^2 y_2 y_1^2 + 1260y_3 y_2^3 y_1 - 945y_2^5.$$

Локализация корня по одному значению функции

9. Рассмотрим случай, когда задан лишь один узел интерполяции x_0. Обычно считают, что если значение $f(x_0)$ «мало», то в окрестности точки x_0 лежит корень уравнения $f(x) = 0$. Конечно, такое заключение нуждается в уточнении, так как, умножая уравнение на малое число, мы можем сделать $f(x_0)$ как угодно малым, не изменив корней уравнения.

Более точные выводы можно получить, используя для оценки малости $f(x_0)$ первую производную $f'(x)$.

Теорема 2.2. Если при некотором $\eta > 0$ функция $f(x)$ непрерывна и дифференцируема на J_x: $[x_0 - \eta, x_0 + \eta]$[1]) и если значение $f(x_0)$ допускает оценку

$$|f(x_0)| \leqslant \eta m, \quad \text{где} \quad m = \inf_{x \in J_x} |f'(x)|,$$

то уравнение $f(x) = 0$ имеет точно один корень на J_x.

Доказательство. Без потери общности можно считать, что $f(x_0) > 0$. По теореме Дарбу производная $f'(x)$ не меняет знака на J_x, поэтому функция $f(x)$ строго монотонна и, значит, принимает наименьшее значение в одном из концов сегмента J_x. Обозначим этот конец буквой ξ. Из неравенства

$$f(\xi) < f(x_0)$$

и из формулы Лагранжа

$$f(x_0) - f(\xi) = (x_0 - \xi) f'(\varrho), \quad \varrho \in J_x,$$

следует, что

$$f(x_0) - f(\xi) = |x_0 - \xi| \cdot |f'(\varrho)| \geqslant \eta m$$

и, значит,

$$f(x_0) - f(\xi) \geqslant f(x_0).$$

Поэтому $f(\xi) \leqslant 0$.

Если $f(\xi) = 0$, то теорема, очевидно, справедлива.

Если $f(\xi) < 0$, то функция $f(x)$ принимает значения разных знаков в точках ξ и x_0 сегмента J_x и в силу непрерывности на J_x имеет корень между точками ξ и x_0. Единственность этого корня на J_x вытекает из монотонности $f(x)$. Теорема доказана.

На практике для локализации корня достаточно проверить выполнение двух неравенств: $|f(x_0)| \leqslant \eta m$, $|f'(x)| \geqslant m$ при $x \in J_x$, т. е. применить в качестве m любую нижнюю грань величины $|f'(x)|$.

Пример. Пусть $|f'(x)| \geqslant 0{,}01$ при $|x - x_0| \leqslant 0{,}001$ и $|f(x_0)| \leqslant 0{,}00001$. Положив $\eta = 0{,}001$, заключаем, что $f(x)$ имеет корень на сегменте $[x_0 - 0{,}001;\ x_0 + 0{,}001]$. Следовательно, x_0 дает приближение к корню с ошибкой, не превосходящей 0,001.

---

[1]) Символ $[a, b]$ мы употребляем для обозначения сегмента (*замкнутого* интервала) с концами a и b, $a \gtrless b$. Символ (a, b) обозначает *открытый* интервал с теми же концами.

ГЛАВА 3

REGULA FALSI И МЕТОД СЕКУЩИХ

Определение regula falsi

1. Пусть $f(x)$ определена на J_x, и пусть *в двух узлах интерполяции* $x_1, x_2 \in J_x$, $x_1 \neq x_2$, заданы значения $f(x_1) = y_1$, $f(x_2) = y_2$, причем $y_1 \neq y_2$.

Для приближенного решения уравнения $f(x) = 0$ мы аппроксимируем функцию $f(x)$ линейной функцией $L(x)$, принимающей значения y_1 и y_2 в точках x_1 и x_2 соответственно:

$$f(x) \sqsupset L(x) = \frac{(x-x_1)\,y_2 - (x-x_2)\,y_1}{x_2 - x_1}. \tag{3.1}$$

Решая линейное уравнение $L(x) = 0$, получаем

$$x = x_3 = \frac{x_1 y_2 - x_2 y_1}{y_2 - y_1}, \tag{3.2}$$

что можно записать также в любой из форм:

$$x_3 = x_1 - y_1 \frac{x_2 - x_1}{y_2 - y_1}, \tag{3.2a}$$

$$x_3 = x_2 - y_2 \frac{x_2 - x_1}{y_2 - y_1}. \tag{3.2b}$$

Формула (3.2) дает *правило ложного положения*, или regula falsi.

2. Исследуем, в какой мере x_3 является приближением к какому-нибудь корню уравнения $f(x) = 0$.

Из (1.5) при $n = 2$ имеем для любого $x \in J_x$:

$$f(x) - L(x) = \frac{1}{2} f''(\xi)(x - x_1)(x - x_2), \quad \xi \in (x, x_1, x_2). \tag{3.3}$$

Здесь (x, x_1, x_2) обозначает *открытый* интервал, концами которого служат крайние из данных трех точек (так что

третья точка лежит внутри интервала). Если $x = x_3$, то $L(x_3) = 0$, и мы получаем

$$f(x_3) = \frac{1}{2} f''(\xi)(x_3 - x_1)(x_3 - x_2), \quad \xi \in (x_1, x_2, x_3). \quad (3.4)$$

Подставляя сюда (3.2a) и (3.2b), преобразуем (3.4) к виду

$$f(x_3) = \frac{1}{2} f''(\xi)\, y_1 y_2 \left(\frac{x_2 - x_1}{y_2 - y_1}\right)^2, \quad \xi \in (x_1, x_2, x_3). \quad (3.5)$$

Но

$$\frac{y_2 - y_1}{x_2 - x_1} = \frac{f(x_2) - f(x_1)}{x_2 - x_1} = f'(\xi_0), \quad \xi_0 \in (x_1, x_2);$$

поэтому из (3.5) окончательно получаем

$$f(x_3) = y_1 y_2 \frac{f''(\xi)}{2f'(\xi_0)^2}, \quad \xi \in (x_1, x_2, x_3), \ \xi_0 \in (x_1, x_2). \quad (3.6)$$

Отсюда видно, что $f(x_3)$ будет мало, если x_1 и x_2 лежат достаточно близко от некоторого корня $f(x)$, так как тогда y_1 и y_2 будут малы.

Применение обратной интерполяции

3. Получим теперь с помощью обратной интерполяции непосредственную оценку отклонения x_3 от корня.

Допустим, что в J_x лежит корень ζ уравнения $f(x) = 0$ и что $f'(x) \neq 0$ на J_x. Тогда для $y = f(x)$ существует обратная функция $x = \varphi(y)$, определенная на y-интервале, соответствующем J_x, и принимающая значение $\varphi(0) = \zeta$. Интерполируя функцию $\varphi(y)$ линейной функцией

$$\chi(y) = \frac{(y - y_1)\, x_2 - (y - y_2)\, x_1}{y_2 - y_1}, \quad (3.7)$$

мы получаем для искомого корня $\zeta = \varphi(0)$ приближенное значение

$$\chi(0) = \frac{x_1 y_2 - x_2 y_1}{y_2 - y_1} = x_3.$$

Для x_3 получено то же значение, что и при прямой интерполяции. Это совпадение объясняется тем, что линейная интерполяция сводится к проведению прямой линии через точки (x_1, y_1) и (x_2, y_2), независимо от того, интерполируем ли мы функцию $f(x)$ или $\varphi(y)$.

4. Для оценки погрешности мы применим формулу (2.2):

$$\varphi(0) - \chi(0) = y_1 y_2 \frac{\varphi''(\eta)}{2}, \quad \eta \in (0, y_1, y_2).$$

Подставим сюда (2.4), полагая $\xi = \varphi(\eta)$, $\xi \in J_x$,

$$\varphi(0) - \chi(0) = -y_1 y_2 \frac{f''(\xi)}{2f'(\xi)^3},$$

или

$$\zeta - x_3 = -y_1 y_2 \frac{f''(\xi)}{2f'(\xi)^3}, \quad \xi \in (\zeta, x_1, x_2). \tag{3.8}$$

Применяя теорему Лагранжа и учитывая, что $f(\zeta) = 0$, получаем

$$y_1 = f(x_1) - f(\zeta) = (x_1 - \zeta) f'(\xi_1), \quad \xi_1 \in (x_1, \zeta),$$
$$y_2 = f(x_2) - f(\zeta) = (x_2 - \zeta) f'(\xi_2), \quad \xi_2 \in (x_2, \zeta)$$

и, окончательно,

$$\zeta - x_3 = \left[-\frac{f''(\xi) f'(\xi_1) f'(\xi_2)}{2f'(\xi)^3} \right] (\zeta - x_1)(\zeta - x_2), \tag{3.9}$$
$$\xi \in (\zeta, x_1, x_2), \quad \xi_1 \in (x_1, \zeta), \quad \xi_2 \in (x_2, \zeta).$$

5. Оценим теперь величину выделенного (квадратными скобками) множителя в (3.9). Пусть всюду в J_x

$$0 < m_1 \leqslant |f'(x)| \leqslant M_1; \quad 0 \leqslant m_2 \leqslant |f''(x)| \leqslant M_2.$$

Тогда модуль выделенного в (3.9) множителя заключен между числами

$$k = \frac{m_2 m_1^2}{2M_1^3} \quad \text{и} \quad K = \frac{M_2 M_1^2}{2m_1^3},$$

и поэтому

$$k|\zeta - x_1||\zeta - x_2| \leqslant |\zeta - x_3| \leqslant K|\zeta - x_1||\zeta - x_2|. \tag{3.10}$$

Если f' и f'' непрерывны в ζ, то при $x_1 \to \zeta$, $x_2 \to \zeta$ формула (3.9) приводит к предельному соотношению.

$$\frac{\zeta - x_3}{(\zeta - x_1)(\zeta - x_2)} \to -\frac{f''(\zeta)}{2f'(\zeta)}. \tag{3.11}$$

Отсюда видно, что если x_1 и x_2 достаточно близки к корню ζ, то x_3 будет давать существенно лучшее приближение к этому корню, чем x_1 и x_2.

Формулу, подобную (3.9), можно получить и из (3.6), заменяя $f(x_3)$ на $f'(\xi')(x_3-\zeta)$, y_1 — на $f'(\xi_1)(x_1-\zeta)$, y_2 — на $f'(\xi_2)(x_2-\zeta)$. Такая замена приводит к формуле

$$\zeta - x_3 = \left[-\frac{f''(\xi)\,f'(\xi_1)\,f'(\xi_2)}{2f'(\xi')\,f'(\xi_0)^2}\right](\zeta - x_1)(\zeta - x_2),$$

в которой первый множитель труднее исследовать, чем в (3.9). Но неравенство (3.10) получается и из этой формулы.

Геометрическое толкование. Условия Фурье

6. Рассмотрим теперь случай, когда график функции $y = f(x)$ подобен изображенному на рис. 1. Взяв x_0 и x_1 за исходные приближения, получим x_2 по regula falsi. Возьмем затем в качестве исходных приближений x_0 и x_2 и получим x_3. Продолжая этот процесс, получим последовательность точек $x_1, x_2, x_3, \ldots$, где

$$x_{\nu+1} = \frac{x_0 f(x_\nu) - x_\nu f(x_0)}{f(x_\nu) - f(x_0)} \quad (\nu = 1, 2, \ldots). \tag{3.12}$$

Сходится ли последовательность $\{x_\nu\}$?

7. Если геометрическая ситуация подобна изображенной на рис. 1, то последовательность $\{x_\nu\}$ сходится. Действительно, точки $x_1, x_2, x_3, \ldots$ лежат со стороны вогнутости

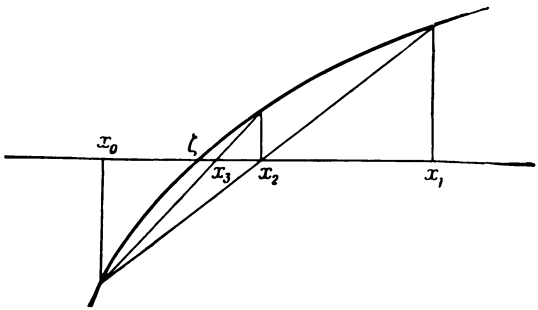

Рис. 1.

кривой и не могут перейти через корень ζ, т. е. они образуют монотонную ограниченную последовательность, которая сходится к некоторому пределу ζ_0.

Совпадает ли ζ_0 с корнем ζ? Вычтем ζ_0 из обеих частей (3.12) и перейдем к пределу при $\nu \to \infty$:

$$0 = \frac{x_0 f(\zeta_0) - \zeta_0 f(x_0) - \zeta_0 [f(\zeta_0) - f(x_0)]}{f(\zeta_0) - f(x_0)} = f(\zeta_0) \frac{x_0 - \zeta_0}{f(\zeta_0) - f(x_0)}.$$

Так как $x_0 \neq \zeta_0$, то отсюда следует, что $f(\zeta_0) = 0$, поэтому ζ_0 — корень $f(x)$ на J_x и, значит, $\zeta_0 = \zeta$.

В проведенном выше доказательстве, помимо принятого ранее условия сохранения знака $f'(x)$ на J_x, предполагалось еще отсутствие точек перегиба графика (достаточным условием этого является сохранение знака $f''(x)$ на J_x). Кроме того, точка x_0 должна быть выбрана так, чтобы выполнялось неравенство $f(x_0) f''(x_0) > 0$. Это так называемые *условия Фурье*.

Если эти условия не будут выполнены, например, если $f(x_0) f''(x_0) < 0$, то мы можем получить расходящуюся последовательность $\{x_\nu\}$. Если последовательность $\{x_\nu\}$ все же будет сходиться к корню, она может быть не монотонной (колебаться около него). Такой случай будет рассмотрен в гл. 5, см. также пример 1 п. 17.

8. Оценка погрешности приближений (3.12) к корню ζ может быть получена из формулы (3.9) после замены в ней x_1, x_2, x_3 на x_0, x_ν, $x_{\nu+1}$:

$$\zeta - x_{\nu+1} = \left[-\frac{f''(\xi) f'(\xi_1) f'(\xi_2)}{2 f'(\xi)^3} (\zeta - x_0) \right] (\zeta - x_\nu), \tag{3.13}$$
$$\xi \in (\zeta, x_0, x_\nu), \quad \xi_1 \in (x_0, \zeta), \quad \xi_2 \in (x_\nu, \zeta).$$

Если выделенный множитель по модулю не превосходит числа ϱ, $\varrho < 1$ и не стремится к нулю с ростом ν, то итерация по формуле (3.12) дает *линейную сходимость:*

$$|\zeta - x_{\nu+1}| \leqslant \varrho |\zeta - x_\nu|.$$

Отсюда получаем

$$|\zeta - x_\nu| \leqslant \varrho^{\nu-1} |\zeta - x_1|, \quad \varrho < 1. \tag{3.14}$$

Формула (3.14) показывает, что при $\varrho = 1/10$ каждый шаг итерации дает дополнительно один верный знак корня, при $\varrho = 1/3$ дополнительный верный знак получается примерно на каждом втором шаге и т. д.

С другой стороны, модуль выделенного множителя будет существенно меньше единицы, если выбрать точку x_0 достаточно близко к корню.

Метод секущих

9. Сходимость метода можно улучшить, если на каждом шаге итерации применять regula falsi к двум последним полученным точкам.

Подставляя в (3.12) вместо *фиксированной* точки x_0 *переменную* точку $x_{\nu-1}$, получаем:

$$x_{\nu+1} = \frac{x_{\nu-1} f(x_\nu) - x_\nu f(x_{\nu-1})}{f(x_\nu) - f(x_{\nu-1})} \quad (\nu = 1, 2, \ldots). \qquad (3.15)$$

Итерацию по формуле (3.15) мы будем в дальнейшем называть *методом секущих*, в отличие от итерации по формуле (3.12). Мы докажем, что метод секущих дает *сверхлинейную сходимость*, если только точки x_0 и x_1 выбраны достаточно близко к корню ζ.

Умножим второе из неравенств (3.10) на K и заменим x_1, x_2, x_3 на $x_{\nu-1}$, x_ν, $x_{\nu+1}$:

$$K\,|\zeta - x_{\nu+1}| \leqslant K^2\,|\zeta - x_{\nu-1}|\,|\zeta - x_\nu|.$$

Полагая $K\,|\zeta - x_\nu| = d_\nu$, получим

$$d_{\nu+1} \leqslant d_{\nu-1}\, d_\nu.$$

Предположим теперь, что d_0 и d_1 не превосходят некоторого числа d, $d < 1$. Тогда

$$d_2 \leqslant d^2,\ d_3 \leqslant d^3,\ d_4 \leqslant d^5,\ d_5 \leqslant d^8, \ldots$$

и вообще

$$d_\nu \leqslant d^{\alpha_\nu},$$

где показатели α_ν образуют последовательность *Фибоначчи*, определяемую условиями

$$\alpha_{\nu+1} = \alpha_\nu + \alpha_{\nu-1} \quad (\nu = 1, 2, 3, \ldots), \quad \alpha_0 = \alpha_1 = 1. \quad (3.16)$$

10. Уравнение (3.16) есть однородное линейное разностное уравнение с постоянными коэффициентами. Мы найдем его общее решение и затем получим частное решение, удовлетворяющее начальным условиям $\alpha_0 = \alpha_1 = 1$. Будем искать решения вида t^ν при $t \neq 0$. Подставляя в уравне-

ния (3.16) $\alpha_\nu = t^\nu$, получаем уравнение $t^{\nu+1} = t^\nu + t^{\nu-1}$, или $t^2 - t - 1 = 0$, корнями которого будут

$$t_1 = \frac{1+\sqrt{5}}{2} \approx 1{,}618; \quad t_2 = \frac{1-\sqrt{5}}{2} \approx -0{,}618. \quad (3.17)$$

В силу линейности и однородности уравнения (3.16) его общим решением будет линейная комбинация

$$\alpha_\nu = c_1 t_1^\nu + c_2 t_2^\nu$$

с произвольными постоянными c_1 и c_2. Выберем эти постоянные так, чтобы удовлетворить начальным условиям $\alpha_0 = \alpha_1 = 1$:

$$c_1 + c_2 = 1; \quad c_1 t_1 + c_2 t_2 = 1.$$

Получим

$$c_1 = \frac{5+\sqrt{5}}{10} = \frac{t_1}{\sqrt{5}}; \quad c_2 = \frac{5-\sqrt{5}}{10} = -\frac{t_2}{\sqrt{5}}.$$

Таким образом, решением задачи (3.16) является последовательность

$$\alpha_\nu = \frac{1}{\sqrt{5}}\left(t_1^{\nu+1} - t_2^{\nu+1}\right) \quad (\nu = 0, 1, 2, \ldots), \qquad (3.18)$$

где t_1, t_2 определены формулами (3.17).

При $\nu \to \infty$ член $\frac{1}{\sqrt{5}} t_1^{\nu+1} \approx 0{,}4472\,(1{,}618)^{\nu+1}$ будет доминировать в (3.18), поэтому имеет место асимптотическая формула:

$$\alpha_\nu \sim \frac{1}{\sqrt{5}} t_1^{\nu+1}.$$

Можно доказать (см. далее гл. 12, п. 11—12), что если последовательность (3.15) сходится (т. е. $\zeta - x_\nu \to 0$), то быстрота этой сходимости определяется соотношением

$$\frac{|\zeta - x_{\nu+1}|}{|\zeta - x_\nu|^{t_1}} \to \left|\frac{2f'(\zeta)}{f''(\zeta)}\right|^{t_2} \quad (\nu \to \infty). \qquad (3.19)$$

Единицы Горнера и индекс эффективности

11. Работу, затраченную на вычисление значения функции или одной из ее производных, мы будем называть *горнером* (*единицей Горнера*). Например, можно сказать, что посредством одного горнера число верных знаков, получае-

мых при вычислении по формуле (3.15), умножается примерно на 1,618 [1]).

Если в процессе решения мы получаем последовательность $\{x_\nu\}$, сходящуюся к ζ, и если для перехода от x_ν к $x_{\nu+1}$ требуется m_ν горнеров, то предел

$$\lim_{\nu\to\infty}\sqrt[m_\nu]{\frac{\ln|x_{\nu+1}-\zeta|}{\ln|x_\nu-\zeta|}},$$

если он существует, мы назовем *индексом эффективности* метода [2]). Из формулы (3.19) мы видим, что индекс эффективности метода секущих [3]) равен $t_1 \approx 1{,}618$.

Правило округления

12. Как следует из (3.8), величина x_3, вычисленная по формуле (3.2), дает аппроксимацию корня ζ с погрешно-

---

[1]) Конечно, употребление единиц Горнера для оценки объема вычислительной работы очень условно. Значимость единиц Горнера меняется от задачи к задаче, а также при увеличении числа десятичных знаков, с которыми ведется расчет.

Однако подобные оценки применяются и в области дифференциальных уравнений с частными производными первого порядка для сравнения различных методов решения. «Индекс простоты» есть обычно сумма порядков всех систем уравнений, применяемых в рассматриваемом методе, хотя очевидно, что конкретные дифференциальные уравнения первого порядка могут оказаться намного сложнее конкретной системы дифференциальных уравнений третьего порядка.

[2]) Предел подкоренного выражения (если он существует) называют иногда *порядком метода*. — *Прим. перев.*

[3]) Причина, по которой regula falsi применяют обычно не в форме (3.15), а в форме (3.12), заключается, очевидно, в известном предубеждении против *экстраполяции* в пользу *интерполяции*; это предубеждение, конечно, оправдано при работе с эмпирическими данными, но гораздо меньше обосновано при работе с аналитическим выражением. (Заметим также, что небольшое изменение алгоритма расчета позволяет вообще избежать экстраполяции. Если известен интервал, отделяющий корень, и если точка

$$x_3 = \frac{x_1 f(x_2) - x_2 f(x_1)}{f(x_2) - f(x_1)}$$

выходит за границы этого интервала, то в дальнейшем вместо точки x_3 можно брать соответствующую границу интервала. — *Прим. перев.*)

стью $O(y_1 y_2)$[1]). Может показаться, что и величины y_1, y_2 должны быть вычислены с такой же степенью точности. Это означало бы, в частности, что величина y_1 должна быть уточнена *после определения порядка величины* y_2. В действительности, однако, при вычислении величины y_1 и y_2 точность порядка $O(y_1 y_2)$ не является необходимой. Мы покажем, что при вычислении y_i необходима и достаточна точность порядка $O[(x_1 - x_2) y_i]$ $(i = 1, 2)$, так что нет надобности в указанном выше уточнении.

13. Обозначим через y_1, y_2 *точные значения* $f(x_1)$, $f(x_2)$, а через $y_1 + \delta_1$, $y_2 + \delta_2$ — их приближенные значения, подставляемые в (3.2). *Погрешность вычисления* x_3 по этим значениям составит $\frac{\partial x_3}{\partial y_1}\delta_1 + \frac{\partial x_3}{\partial y_2}\delta_2$. Для того чтобы эта погрешность имела порядок $O(y_1 y_2)$, погрешности δ_1 и δ_2 должны удовлетворять условиям

$$\frac{\partial x_3}{\partial y_1}\delta_1 \equiv y_2 \frac{x_1 - x_2}{(y_1 - y_2)^2}\delta_1 = O(y_1 y_2),$$

$$\frac{\partial x_3}{\partial y_2}\delta_2 \equiv -y_1 \frac{x_1 - x_2}{(y_1 - y_2)^2}\delta_2 = O(y_1 y_2),$$

откуда

$$\delta_i = O\left[y_i \frac{(y_1 - y_2)^2}{x_1 - x_2}\right] \quad (i = 1, 2).$$

При $x_1 \to \zeta$, $x_2 \to \zeta$ имеем $y_1 - y_2 \sim f'(\zeta)(x_1 - x_2)$, и указанные выше условия дают

$$\delta_i = O[(x_1 - x_2) y_i] \quad (i = 1, 2), \text{ ч. т. д.} \tag{3.20}$$

Условие (3.20) может быть также записано в виде

$$\delta_i = O(y_i^2 - y_1 y_2). \tag{3.21}$$

---

[1]) Мы пишем $f = O(g)$ для некоторого предельного процесса, если $\overline{\lim}\left|\frac{f}{g}\right|$ конечен (в частности, если совокупность значений $\left|\frac{f}{g}\right|$ ограничена); запись $f = o(g)$ означает, что $\lim \frac{f}{g} = 0$ для данного предельного процесса. Конечно, при употреблении этих формул рассматриваемый предельный процесс должен быть точно определен.

Если для фиксированного положительного ε имеет место неравенство $\left|\frac{y_2}{y_1}\right| \leqslant 1-\varepsilon$, то условия (3.21) эквивалентны условиям

$$\delta_1 = O(y_1^2),\ \delta_2 = O(y_1 y_2)\ \ (|y_2| \leqslant (1-\varepsilon)|y_1|). \quad (3.22)$$

Наконец, если $y_1 y_2 < 0$, то условия (3.22) являются достаточными при любой величине отношения $\left|\frac{y_2}{y_1}\right|$.

14. Выше мы установили связь только между погрешностями исходных данных (y_1, y_2) и обусловленной ими погрешностью формулы (3.2). Вопрос о том, какая из теоретически эквивалентных формул (3.2), (3.2a), (3.2b) дает меньшую погрешность при вычислениях, требует отдельного рассмотрения.

При расчете по формуле (3.2) числитель должен быть вычислен с точностью $O[y_1 y_2 (y_1 - y_2)] = O[y_1 y_2 (x_1 - x_2)]$. Положение существенно улучшается при вычислениях по одной из формул (3.2a) или (3.2b) [1]). В частности, применяя формулу (3.2b) при условии $|y_2| \leqslant |y_1|$, мы можем при вычислении произведения $y_2 \frac{x_2 - x_1}{y_2 - y_1}$ ограничиться точностью $O(y_1 y_2)$. Поэтому отношение $\frac{x_2 - x_1}{y_2 - y_1}$ $\left(\text{которое стремится к } \frac{1}{f'(\zeta)}\right)$ достаточно вычислить только с точностью $O(y_1)$.

Локализация корня с помощью regula falsi

15. До сих пор мы предполагали, что в рассматриваемом интервале лежит корень, а x_0 и x_1 достаточно близки к корню, чтобы d_0 и d_1 были меньше *единицы*. Однако часто мы не знаем заранее, лежит ли корень в рассматриваемом интервале, и хотим в процессе вычислений получить ответ и на этот вопрос.

Мы уже показали в теореме 2.2, что для локализации корня уравнения $f(x) = 0$ в окрестности некоторой точки x_*

---

[1]) То есть при *вычислении поправок* $x_3 - x_1$ или $x_3 - x_2$ к значениям x_1 или x_2. — *Прим. перев.*

надо, чтобы модуль $|f(x_*)|$ был достаточно мал; точнее, корень будет лежать на сегменте $[x_* - \eta, x_* + \eta]$, если $|f(x_*)| \leqslant \eta m$, где $m = \inf |f'(x)|$ на этом сегменте.

Пусть теперь по начальным точкам x_1 и x_2 найдено значение x_3 методом regula falsi (3.2). Нас интересует вопрос, существует ли корень уравнения $f(x) = 0$ в окрестности точки x_3. Из (3.6) имеем

$$f(x_3) = \frac{f''(\xi)}{2f'(\xi_0)^2} f(x_1) f(x_2), \quad \xi \in (x_1, x_2, x_3), \quad \xi_0 \in (x_1, x_2).$$

Если известны величины

$$m_1 = \inf_{x \in (x_1, x_2)} |f'(x)|, \quad M_2 = \sup_{x \in (x_1, x_2, x_3)} |f''(x)|,$$

то можно избежать вычисления $f(x_3)$. Мы должны лишь проверить, выполняется ли при некотором η неравенство

$$\frac{M_2}{2m_1^2} |f(x_1)| |f(x_2)| \leqslant \eta m, \tag{3.23}$$

где $m = \inf |f'(x)|$ на сегменте $[x_3 - \eta, x_3 + \eta]$ [1]).

16. Мы проиллюстрируем сказанное выше на примере классического уравнения, впервые рассмотренного Ньютоном:

$$f(x) \equiv x^3 - 2x - 5 = 0;$$

$$x_1 = 1; \quad y_1 = -6; \quad x_2 = 2; \quad y_2 = -1; \quad x_3 = 2{,}2.$$

Производные $f'(x) = 3x^2 - 2$, $f''(x) = 6x$ монотонно возрастают в интервалах (x_1, x_2) и (x_1, x_2, x_3) соответственно. Поэтому $f'(x)$ имеет наименьшее значение на левом конце

---

[1]) Более простой и точный способ локализации корня мы получим, применяя вместо формулы (3.6) формулу (3.4):

$$f(x_3) = \frac{f''(\xi)}{2} (x_3 - x_1)(x_3 - x_2).$$

При этом достаточно проверить выполнение неравенства

$$\frac{M_2}{2} |x_3 - x_1| |x_3 - x_2| \leqslant \eta m, \tag{3.23a}$$

которое в отличие от неравенства (3.23) не содержит m_1. — *Прим. перев.*

(x_1, x_2), а $f''(x)$ имеет наибольшее значение на правом конце (x_1, x_2, x_3). Подставляя эти значения в (3.23), получим

$$\left(\frac{13,2}{2 \cdot 1^2}\right) \cdot 6 = 39,6 \leqslant \eta m.$$

Здесь при любом $\eta \leqslant 1$ имеем $m < f'(x_3) = 12,52$, и мы не можем гарантировать существование корня на $[x_3 - 1,\ x_3 + 1]$[1]).

Возьмем теперь другие начальные приближения: $x_1 = 1,8$; $y_1 = -2,768$; $x_2 = 2$; $y_2 = -1$; $x_3 = 2,113$.

Подставляя новое приближение в (3.23), получим

$$\frac{12,678}{2\,(7,72)^2} \cdot 2,768 = 0,2944 \leqslant \eta m.$$

Так как вблизи точки x_3 производная $f' > 10$, то пробуем $\eta = 0,03$; при этом $m = f'(2,083) > 11,016$, $0,2944 < 0,03 \cdot 11$, и мы можем утверждать, что x_3 дает приближение к корню с погрешностью, не превосходящей 0,03, т. е. что в интервале $(2,113 \pm 0,03)$ существует корень.

Примеры вычислений по regula falsi

17. Перепишем формулы (3.12) и (3.15) в форме, удобной для расчетов:

$$x_{\nu+1} = x_\nu - f(x_\nu)\,\frac{x_\nu - x_0}{f(x_\nu) - f(x_0)}\,; \qquad (3.24а)$$

---

[1]) Оценка (3.23а), приведенная в сноске на стр. 31, позволяет локализовать корень даже в этом случае, ибо неравенству

$$\frac{13,2}{2} \cdot 1,2 \cdot 0,2 = 1,584 \leqslant \eta m$$

удовлетворяет $\eta = 0,15$ ($m = f'(2,05) > 10,6$; $1,584 < 0,15 \cdot 10,6$) Таким образом, корень лежит в интервале $(2,2 \pm 0,15)$.

Для начальных приближений $x_1 = 1,8$, $x_2 = 2$ та же оценка дает неравенство

$$\frac{12,678}{2} \cdot 0,313 \cdot 0,113 = 0,225 \leqslant \eta m,$$

которому удовлетворяет $\eta = 0,021$ ($m = f'(2,092) > 11,12$); это локализует корень в интервале $(2,113 \pm 0,021)$. — *Прим. перев.*

$$x_{\nu+1} = x_\nu - f(x_\nu)\frac{x_\nu - x_{\nu-1}}{f(x_\nu) - f(x_{\nu-1})}. \qquad (3.24b)$$

Если с помощью этих формул мы хотим получить k знаков после запятой в $x_{\nu+1}$, а $|f(x_\nu)| < 10^{-k_1}$, то значение дроби достаточно вычислить с $(k - k_1)$ знаками после запятой.

Пример 1. Вычисление корня уравнения

$$f(x) \equiv x^3 - 2x - 5 = 0$$

по формуле (3.24a), дающей линейную сходимость.

Значение корня с 13 десятичными знаками:

$$\zeta = 2{,}0945514815423.$$

ν	x_ν	$\zeta - x_\nu$
0	2*	
1	3	$-0{,}9054$
2	2,0588235294	$+0{,}0357$
3	2,0965586362	$-0{,}00201$
4	2,0944405193	$+0{,}0_3111$**
5	2,0945576218	$-0{,}0_5614$
6	2,0945511399	$+0{,}0_6342$
7	2,0945515006	$-0{,}0_7191$

* Здесь условия Фурье для точки $x_0 = 2$ не выполнены, так как $f'' > 0$, $f(2) < 0$. Поэтому приближения x_ν колеблются около корня *Прим. перев.*

** Индекс указывает число нулей после запятой.

Пример 2. (Метод секущих). Вычисление того же корня по формуле (3.24b), дающей сверхлинейную сходимость.

Значение корня, верное до 24 знака, есть

$$\zeta = 2{,}094551481542326591482387.$$

ν	x_ν	$\zeta - x_\nu$	$\dfrac{\zeta - x_{\nu-1}}{\zeta - x_\nu}$
0	2		
1	3	—0,9054	
2	2,0588235294	0,0357	0,411
3	2,08126365965	0,0133	0,574
4	2,09482418427	$-0,0_{3}273$	0,565
5	2,09454943175	$0,0_{5}205$	0,5624
6	2,094551481228	$0,0_{9}314$	0,5629779
7	2,0945514815423269542	$-0,0_{15}363$	*

* Вычисления для x_7 проводились с удвоенной точностью, т. е. до 20 знака.

Здесь

$$f'(\zeta) = 11,1614377, \quad f''(\zeta) = 12,56730888,$$

$$\frac{f''(\zeta)}{2f'(\zeta)} = 0,5629789. \tag{3.25}$$

Дальнейшее исследование regula falsi с другой точки зрения дано в гл. 5.

Обобщение regula falsi на случай системы двух уравнений дано в приложении IV.

ГЛАВА 4

ИТЕРАЦИЯ

Критерий сходимости итерации

1. Пусть $\psi(x)$ определена на J_x. Выберем $x_1 \in J_x$ и построим

$$x_2 = \psi(x_1), \quad x_3 = \psi(x_2), \quad \ldots, \quad x_{\nu+1} = \psi(x_\nu), \quad \ldots \tag{4.1}$$

Мы предполагаем, что все точки последовательности $\{x_\nu\}$ принадлежат J_x. В частности, если $\psi(x_1) = x_1$, то все члены последовательности $\{x_\nu\}$ совпадают с x_1.

Число ζ, для которого $\psi(\zeta) = \zeta$, называется *неподвижной точкой*, или *центром итерации* (4.1); функция $\psi(x)$ называется *итерирующей функцией*.

Лемма 4.1. Пусть $F(x)$ определена на J_x и

$$F(x) \equiv x - \psi(x); \tag{4.2}$$

возьмем $\psi(x)$ в качестве итерирующей функции для построения последовательности (4.1). Если $x_\nu \to \zeta$ и $\psi(x)$ непрерывна в ζ, то $F(\zeta) = 0$.

Доказательство. Переходя к пределу в равенстве $x_{\nu+1} = \psi(x_\nu)$ и учитывая непрерывность $\psi(x)$ в ζ, получим $\zeta = \psi(\zeta)$, или $F(\zeta) = \zeta - \psi(\zeta) = 0$, ч. т. д.

Точки притяжения и отталкивания

2. Через $U(\zeta_0)$ мы обозначаем симметричную окрестность точки ζ_0, т. е. интервал $(\zeta_0 - \eta, \zeta_0 + \eta)$. Пусть $\zeta_0 = \psi(\zeta_0)$. Мы будем называть ζ_0 *точкой притяжения*, если существует такая ее окрестность $U(\zeta_0)$, что последовательность (4.1) сходится к ζ_0 при любой начальной точке x_1 из этой окрестности. Мы будем называть ζ_0 *точкой отталкивания*, если существует такая окрестность $U(\zeta_0)$, что последовательность (4.1) не сходится к ζ_0, с какой бы точки

$x_1 \in U(\zeta_0)$ мы ни начинали (кроме тех случаев, когда одна из точек x_ν совпадает с точкой ζ_0)[1]).

Лемма 4.2. Пусть итерирующая функция $\psi(x)$ определена на J_x. Предположим, что в (J_x) имеется неподвижная точка ζ_0: $\zeta_0 = \psi(\zeta_0)$ и что существует производная $\psi'(\zeta_0)$. При этом точка ζ_0 будет точкой притяжения, если $|\psi'(\zeta_0)| < 1$, и точкой отталкивания, если $|\psi'(\zeta_0)| > 1$. В первом случае

$$\frac{x_{\nu+1} - \zeta_0}{x_\nu - \zeta_0} \to \psi'(\zeta_0). \tag{4.3}$$

3. Доказательство. *Часть I.* $|\psi'(\zeta_0)| < 1$. Выберем число p так, чтобы $|\psi'(\zeta_0)| < p < 1$. Тогда найдется такая окрестность $U(\zeta_0)$, в которой

$$\left|\frac{\psi(x) - \psi(\zeta_0)}{x - \zeta_0}\right| \leqslant p. \tag{4.4}$$

Возьмем любую начальную точку $x_1 \in U(\zeta_0)$. Тогда

$$\left|\frac{x_2 - \zeta_0}{x_1 - \zeta_0}\right| \leqslant p < 1;$$

следовательно, точка x_2 лежит к ζ_0 ближе, чем x_1, и $x_2 \in U(\zeta_0)$. Применяя неравенство (4.4) последовательно к $x_2, \ldots, x_\nu$, получим $|x_3 - \zeta_0| \leqslant p\,|x_2 - \zeta_0| \leqslant p^2 |x_1 - \zeta_0|, \ldots,$

$$|x_{\nu+1} - \zeta_0| \leqslant p^\nu\,|x_1 - \zeta_0| \to 0 \qquad (\nu \to \infty). \tag{4.5}$$

Таким образом, ζ_0 — точка притяжения, $x_\nu \to \zeta_0$,

$$\frac{x_{\nu+1} - \zeta_0}{x_\nu - \zeta_0} = \frac{\psi(x_\nu) - \psi(\zeta_0)}{x_\nu - \zeta_0} \to \psi'(\zeta_0).$$

4. Доказательство. *Часть II.* $|\psi'(\zeta_0)| > 1$. Выберем число p так, чтобы $|\psi'(\zeta_0)| > p > 1$. Тогда найдется такая окрестность $U(\zeta_0)$, в которой

$$\left|\frac{\psi(x) - \psi(\zeta_0)}{x - \zeta_0}\right| \geqslant p. \tag{4.5a}$$

---

[1]) Подобным образом можно ввести и понятия *точек притяжения* и *отталкивания* при *односторонней* аппроксимации, заменяя окрестность $U(\zeta_0)$ на соответствующую *одностороннюю* окрестность.

Для любой начальной точки $x_1 \in U(\zeta_0)$ будем иметь $|x_2 - \zeta_0| \geqslant p\,|x_1 - \zeta_0|$, поэтому точка x_2 лежит дальше от ζ_0, чем x_1. Это же утверждение верно и для любой точки x_ν, если x_ν остается в $U(\zeta_0)$ и $x_\nu \neq \zeta_0$. Таким образом, x_ν не стремится к ζ_0, если $x_\nu \neq \zeta_0$ ни при каком ν; ζ_0 есть точка отталкивания.

5. Замечания. (a) Вне окрестности $U(\zeta_0)$ неравенство (4.5a) может не выполняться, поэтому если одна из точек x_ν выйдет за окрестность $U(\zeta_0)$, то какая-либо из последующих точек может оказаться равной ζ_0. Можно построить неаналитическую функцию $\psi(x)$, равную ζ_0 во всех точках вне определенной окрестности точки ζ_0. Можно даже построить аналитическую функцию $\psi(x)$, равную ζ_0 на счетном множестве точек вне $U(\zeta_0)$.

(b) Если $|\psi'(\zeta_0)| < 1$, то сходимость будет тем лучше, чем меньше $|\psi'(\zeta_0)|$ (конечно, при условии выбора начальной точки в достаточно малой окрестности точки ζ_0). Вообще, быстрота сходимости зависит от того, насколько мал модуль отношения

$$\left|\frac{\psi(x) - \psi(\zeta_0)}{x - \zeta_0}\right|.$$

(c) Так как нашей целью является нахождение ζ_0, нам важно знать, будет ли ζ_0 точкой притяжения или точкой отталкивания. Если во всем рассматриваемом интервале $|\psi'(x)| \leqslant p < 1$ и если в этом интервале лежит корень ζ_0, то точка ζ_0 является точкой притяжения и может быть найдена посредством итерации. В этом случае справедливо неравенство (4.5); сходимость будет быстрой, если p достаточно мало. На практике сходимость называют «хорошей», если $p < \frac{1}{10}$, и «медленной», если $p > \frac{1}{2}$. Мы дадим методы улучшения сходимости для любого случая.

Улучшение сходимости

6. Пусть $|\psi'(x)| < 1$ в рассматриваемом интервале. Что означает это условие для $F(x) = x - \psi(x)$? Мы имеем $\psi'(x) = 1 - F'(x)$ и, значит,

$$|1 - F'(x)| < 1, \quad \text{т. е.} \quad 0 < F'(x) < 2.$$

Таким образом, опираясь на полученный выше критерий сходимости, мы можем применять метод итерации непосредственно лишь к таким функциям $F(x)$, которые монотонно *возрастают* и к тому же не слишком быстро. [Если $F(x)$ монотонно *убывает*, мы заменим $F(x)$ на $-F(x)$.]

Допустим, что $F'(x)$ *положительна* в рассматриваемом интервале, $0 < m \leqslant F'(x) \leqslant M$. Корни уравнения $F(x) = 0$ не изменятся и $F'(x)$ останется положительной, если мы умножим $F(x)$ на постоянную $c > 0$. Постоянную c можно выбрать так, чтобы обеспечить сходимость даже при $M > 2$ и улучшить ее в общем случае. Положим $F_c(x) = cF(x)$ и $\psi_c(x) = x - cF(x)$. Тогда производная $\psi_c'(x) = 1 - cF'(x)$ заключена между $1 - cM$ и $1 - cm$, так что $|\psi_c'(x)| \leqslant \max(|1 - cM|, |1 - cm|)$.

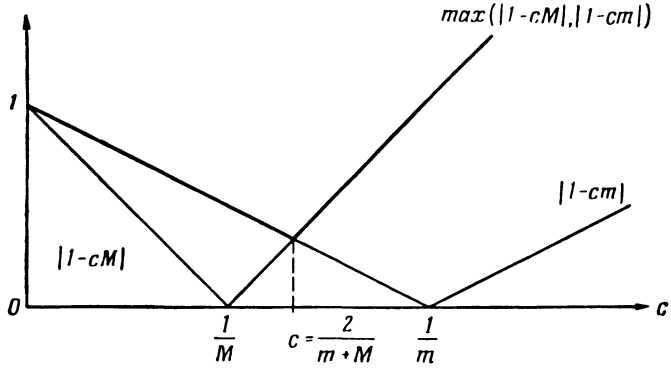

Рис. 2.

Как видно из рис. 2, величина $\max(|1 - cM|, |1 - cm|)$ достигает наименьшего значения, когда

$$1 - cm = -(1 - cM),$$

т. е. при

$$c = \frac{2}{m + M}. \tag{4.6}$$

Это приводит к оценке

$$|\psi_c'(x)| \leqslant \frac{M-m}{M+m}. \tag{4.7}$$

Отношение $\frac{M-m}{M+m}$ будет довольно мало, если только $F'(x)$ изменяется не слишком быстро, т. е. если $|F''(x)|$ мал. Во всяком случае, если $F''(x)$ ограничена в окрестности корня, мы можем сделать верхнюю грань в (4.7) сколь угодно малой и, значит, сходимость сколь угодно быстрой за счет соответствующего сужения рассматриваемой окрестности корня.

В отдельных случаях из предварительных теоретических исследований бывает известна величина $\psi'(\zeta_0) = \alpha$. Если $\alpha \neq 0$, $\alpha \neq 1$, то можно добиться заметного улучшения сходимости, заменив $\psi(x)$ на

$$\psi_*(x) = \frac{1}{1-\alpha}[\psi(x) - \alpha x]. \tag{4.8}$$

Действительно, ζ_0 является неподвижной точкой также и для $\psi_*(x)$, а производная $\psi_*'(\zeta_0)$ равна нулю, как видно из равенств

$$\psi_*(\zeta_0) = \frac{1}{1-\alpha}[\psi(\zeta_0) - \alpha\zeta_0] = \zeta_0,$$

$$\psi_*'(\zeta_0) = \frac{1}{1-\alpha}[\psi'(\zeta_0) - \alpha] = 0.$$

Если величина $\psi'(\zeta_0)$ неизвестна, можно все же попытаться использовать выражение (4.8), взяв вместо $\psi'(\zeta_0)$ какое-либо ее приближенное значение. Например, если $\psi'(x)$ непрерывна в ζ_0, то при $x_0 \to \zeta_0$, $x_1 \to \zeta_0$ имеем

$$\frac{x_2 - x_1}{x_1 - x_0} = \frac{\psi(x_1) - \psi(x_0)}{x_1 - x_0} \to \alpha;$$

поэтому можно в качестве приближенного значения α взять $\frac{x_2 - x_1}{x_1 - x_0}$. Заменим в (4.8) α на $\frac{x_2 - x_1}{x_1 - x_0}$:

$$\frac{x_1 - x_0}{x_2 - 2x_1 + x_0}\left[\frac{x_2 - x_1}{x_1 - x_0}x - \psi(x)\right].$$

Положив здесь $x = x_1$, получим новую точку

$$x^* = \frac{x_0 x_2 - x_1^2}{x_2 - 2x_1 + x_0} = \frac{x_0\psi[\psi(x_0)] - [\psi(x_0)]^2}{\psi[\psi(x_0)] - 2\psi(x_0) + x_0} \equiv \Psi(x_0),$$

которая в большинстве случаев существенно ближе к корню ζ_0, чем x_2. Итерирующая функция $\Psi(x)$ была введена Стеффенсеном, который получил ее другим путем (см. приложение V).

7. Проиллюстрируем изложенный метод улучшения сходимости на примере уравнения

$$x + \lg x = 0{,}5,$$

которое Уиттекер и Робинсон [1]) решали с помощью итерации

$$x_{\nu+1} = 0{,}5 - \lg x_\nu,$$

взяв за начальное значение $x_0 = 0{,}68$. Для ускорения сходимости они брали на каждом шаге среднее арифметическое $\frac{1}{2}(x_\nu + x_{\nu+1})$ и за 7 шагов расчета получили приближенное значение корня $x=0{,}672382$, верное лишь до пятого знака.

На сегменте [0,67; 0,68] функция $F(x) = x + \lg x - 0{,}5$ имеет первую производную $F'(x) = 1 + \frac{\lg e}{x}$ ($\lg e = 0{,}43429448$), заключенную между $m = 1{,}63866$ и $M = 1{,}64821$. По формуле (4.6) определяем $c = 0{,}6085$ и строим $F_c(x) = cF(x)$;

$$\psi_c(x) = x - cF(x) = 0{,}3915\,x - 0{,}6085 \lg x + 0{,}30425.$$

При этом $|\psi_c'(x)| < 0{,}0029$ на сегменте [0,67; 0,68].

Две итерации по формуле $x_{\nu+1} = \psi_c(x_\nu)$ дают

$$x_0 = 0{,}68; \quad x_1 = 0{,}67239; \quad x_2 = 0{,}672383185.$$

Для оценки погрешности приближения x_2 вычислим $F_c(x_2) = (21{,}2 \pm 0{,}6) \cdot 10^{-9}$ и применим теорему 2.2. Так как $|F_c(x_2)| < 21{,}8 \cdot 10^{-9}$, а $F_c'(x) = 1 - \psi_c'(x) > 1 - 0{,}0029$, то корень ζ_0 лежит в интервале $(x_2 \pm 22 \cdot 10^{-9})$. Таким образом, за два шага расчета мы получаем приближение, погрешность которого не превосходит $22 \cdot 10^{-9} < 3 \cdot 10^{-8}$.

8. К формуле (4.6) для величины c мы приходим с помощью следующих рассуждений. Мы выбираем c так, чтобы значение $|\psi_c'(\zeta_0)| = |1 - cF'(\zeta_0)|$ было как можно меньше. При

[1]) Уиттекер Э. Т., Робинсон Д., Математическая обработка результатов наблюдений. М., ОНТИ, 1935, § 43, пример 2.

этом наиболее подходящим значением для $\frac{1}{c}$ оказывается *среднее арифметическое* между наименьшим (m) и наибольшим (M) значениями F'. На практике обычно делят $F(x)$ на значение производной $F'(\xi)$, выбирая точку ξ соответствующим образом, и получают сходимость во всех случаях, когда $F'(\xi) > \frac{1}{2} M$.

Конечно, если мы заранее знаем, что x_ν намного ближе к ζ_0, чем $x_{\nu-1}$, то лучше всего брать на ν-м шаге $c_\nu = \frac{1}{F'(x_\nu)}$, т. е. вычислять $x_{\nu+1}$ по формуле

$$x_{\nu+1} = x_\nu - \frac{F(x_\nu)}{F'(x_\nu)}.$$

Таким путем мы приходим к формуле Ньютона [1]).

До сих пор в наших исследованиях методов итерации было очень существенным предположение о том, что $F(x)$ монотонно возрастает (или монотонно убывает). В следующей главе мы рассмотрим метод, дающий результат даже при изменении знака $F'(x)$ в рассматриваемом интервале, но при условии, что модуль $|F'(x)|$ остается ограниченным.

---

[1]) В оригинале эта формула всюду называется формулой Ньютона — Рафсона. Мы сохраняем название, установившееся в русской литературе. — *Прим. перев.*

ГЛАВА 5

ДАЛЬНЕЙШЕЕ ИССЛЕДОВАНИЕ ИТЕРАЦИЙ. КРАТНЫЕ КОРНИ

Итерации посредством монотонных функций

1. Теорема 5.1. Пусть $f(x)$ непрерывна на J_0: $[x_0, x_0+d]$, причем d выбрано так, что $f(x_0)\,d<0$; пусть

$$\frac{f(x')-f(x)}{x'-x}\leqslant 1 \quad (x, x'\in J_0). \tag{5.1}$$

Положим $\psi(x)=x-f(x)$ и будем строить x_ν $(\nu=0, 1, \ldots)$ посредством итерации $x_{\nu+1}=\psi(x_\nu)$ до тех пор, пока $x_\nu\in J_0$. Тогда:

a) если $f(\zeta)=0$, $\zeta\in J_0$, и $f(x)\neq 0$ в (x_0, ζ), то

$$|x_\nu-\zeta|\downarrow 0 \quad \text{при} \quad \nu\to\infty \text{ и все } x_\nu\in J_0\ ^{1)};$$

b) если $f(x)\neq 0$ в J_0, то существует такое число n_0, что x_{n_0} не лежит в J_0.

2. Доказательство. Покажем, прежде всего, что функция $\psi(x)$ монотонно *возрастает* в J_0. Действительно,

$$\frac{\psi(x')-\psi(x)}{x'-x}=\frac{x'-f(x')-[x-f(x)]}{x'-x}=1-\frac{f(x')-f(x)}{x'-x}\geqslant 0$$

и, значит, $\psi(x')-\psi(x)\geqslant 0$ при $x'>x$.

Пусть теперь $x_1=\psi(x_0)=x_0-f(x_0)$.

Если $d>0$, то $f(x_0)<0$ и $x_0<x_1$.

Если $d<0$, то $f(x_0)>0$ и $x_0>x_1$.

Дальнейшие итерации посредством возрастающей функции $\psi(x)$ приводят к монотонной последовательности $\{x_\nu\}$. Действительно, в случае $d>0$ из $x_0<x_1$ следует $\psi(x_0)\leqslant\psi(x_1)$, $x_1\leqslant x_2$, далее $\psi(x_1)\leqslant\psi(x_2)$, $x_2\leqslant x_3$, ..., и мы получаем *возрастающую* последовательность $x_0<x_1\leqslant x_2\leqslant x_3\leqslant\ldots$. В случае $d<0$ из $x_0>x_1$ следует

---

[1]) Мы пишем $x\downarrow\zeta$, если x стремится к ζ, монотонно *убывая* справа); $x\uparrow\zeta$ означает, что x стремится к ζ, монотонно *возрастая* (слева).

$\psi(x_0) \geqslant \psi(x_1)$, $x_1 \geqslant x_2$. Далее, $\psi(x_1) \geqslant \psi(x_2)$, $x_2 \geqslant x_3$, ..., и мы получаем *убывающую* последовательность $x_0 > x_1 \geqslant x_2 \geqslant \geqslant x_3 \geqslant \ldots$.

Если на ν-м шаге имеет место равенство

$$x_{\nu+1} = x_\nu - f(x_\nu) = x_\nu,$$

то

$$f(x_\nu) = 0$$

и последовательность сходится к x_ν ($x_\nu = x_{\nu+1} = x_{\nu+2} = \ldots$).

3. *Если все x_ν лежат на сегменте J_0*, то монотонная последовательность $\{x_\nu\}$ ограничена и поэтому имеет предел $\zeta_0 \in J_0$. В силу леммы 4.1 имеем

$$f(\zeta_0) = 0.$$

Этим доказано утверждение (b).

Пусть теперь выполнены условия (a). Допустим, что $d > 0$. Тогда $x_1 > x_0$ и $x_1 = \psi(x_0) \leqslant \psi(\zeta) = \zeta$, т. е. x_1 лежит *между* x_0 и ζ. Повторяя эти рассуждения, заключаем, что все x_ν лежат между x_0 и ζ, образуя возрастающую последовательность. Ее предел есть корень уравнения $f(x) = 0$, а так как ζ — ближайший к x_0 корень этого уравнения, то $x_\nu \uparrow \zeta$. В случае $d < 0$ аналогичные рассуждения дают $x_\nu \downarrow \zeta$. Теорема полностью доказана.

4. *Если $f(x)$ имеет конечную первую производную на J_0, то условие* (5.1) *можно заменить на*

$$f'(x) \leqslant 1 \qquad (x \in J_0). \tag{5.2}$$

Действительно, (5.2) получается из (5.1) переходом к пределу при $x' \to x$, а (5.1) получается из (5.2) с помощью теоремы Лагранжа

$$\frac{f(x') - f(x)}{x' - x} = f'(\xi), \quad \xi \in (x', x). \tag{5.3}$$

5. Замечание. Если не выполнено (5.2) или условие $f(x_0)\,d < 0$, то вместо уравнения $f(x) = 0$ можно рассматривать уравнение $cf(x) = 0$, где c выбирается так, чтобы выполнялись неравенства:

$$cf'(x) \leqslant 1 \quad \text{и} \quad cf(x_0)\,d < 0.$$

В гл. 4 было показано, что можно добиться улучшения сходимости за счет выбора подходящего множителя c в уравнении $cf(x) = 0$. Однако рассмотренный там метод не обеспечивает сходимости в случае $f'(\zeta_0) = 0$, т. е. в случае кратного корня. Если же выполнены условия теоремы 5.1, то при $\psi'(\zeta_0) = 1$ мы еще получаем сходимость, хотя и очень слабую. Рассмотрим теперь этот случай подробнее

Кратные корни

6. Пусть ζ — корень $f(x)$ точной кратности k, и пусть $f(x)$ k раз дифференцируема в точке ζ; тогда

$$f(\zeta) = 0, \quad f'(\zeta) = 0, \ \ldots, \ f^{k-1}(\zeta) = 0; \quad f^{(k)}(\zeta) \neq 0 \tag{5.4}$$

и при $x \to \zeta$ имеем [1])

$$\frac{f(x)}{(x-\zeta)^k} \to \frac{f^{(k)}(\zeta)}{k!} \equiv A_0 \neq 0. \tag{5.5}$$

Для дальнейшего исследования мы ограничимся предположением, что $f(x)$ в точке ζ обращается в нуль таким образом, что для некоторого положительного α существует конечный предел

$$\lim \frac{f(x)}{|x-\zeta|^{1+\alpha}} = A \neq 0 \tag{5.6}$$

(или при $x \uparrow \zeta$, или при $x \downarrow \zeta$).

Для нас несущественно, будет ли α целым или нет. По теореме 5.1 последовательность $\{x_\nu\}$, определенная итерацией

$$x_{\nu+1} = x_\nu - f(x_\nu), \tag{5.7}$$

при подходящем начальном значении x_0 монотонно стремится к ζ. Мы исследуем теперь быстроту этой сходимости.

---

[1]) Формула (5.5) может быть получена, например, $(k-1)$-кратным применением правила Лопиталя:

$$\lim_{x\to\zeta} \frac{f(x)}{(x-\zeta)^k} = \lim_{x\to\zeta} \frac{f(x)-f(\zeta)}{(x-\zeta)^k} = \lim_{x\to\zeta} \frac{f'(x)-f'(\zeta)}{k(x-\zeta)^{k-1}} = \ldots =$$

$$= \lim_{x\to\zeta} \frac{f^{(k-1)}(x) - f^{(k-1)}(\zeta)}{k!(x-\zeta)} = \frac{f^{(k)}(\zeta)}{k!}.$$ — *Прим. перев.*

Теорема 5.2. Пусть $f(x)$ определена в односторонней окрестности J точки ζ: $[\zeta, \zeta+\delta]$ и односторонне непрерывна в ζ, причем $f(\zeta)=0$. Пусть для некоторого $\alpha>0$ существует конечный предел (5.6) при $x\to\zeta$, $x\in J$. Если при этих условиях последовательность x_ν, определенная итерацией (5.7), лежит на J и сходится к ζ, то

$$\nu^{1/\alpha}|x_\nu-\zeta|\to\frac{1}{|\alpha A|^{1/\alpha}}\quad(\nu\to\infty). \tag{5.8}$$

7. Доказательство. При $x_\nu\to\zeta$ из (5.6) следует, что

$$f(x_\nu)\sim A|x_\nu-\zeta|^{1+\alpha};$$

отсюда видно, что при достаточно больших ν $(\nu>n_0)$ значения $f(x_\nu)$, а следовательно, и отношения $\frac{f(x_\nu)}{x_\nu-\zeta}$ сохраняют знак (напомним, что $x_\nu\in J$, а ζ — конец сегмента J). Более того, из тождества

$$\frac{x_{\nu+1}-\zeta}{x_\nu-\zeta}=\frac{x_\nu-f(x_\nu)-\zeta}{x_\nu-\zeta}=1-\frac{f(x_\nu)}{x_\nu-\zeta} \tag{5.9}$$

следует, что эти отношения положительны:

$$\frac{f(x_\nu)}{x_\nu-\zeta}>0\quad(\nu>n_0)$$

(в противном случае отношения $(x_{\nu+1}-\zeta):(x_\nu-\zeta)$ были бы больше 1 и x_ν не могли бы стремиться к ζ). Поэтому

$$\frac{f(x_\nu)}{x_\nu-\zeta}\sim|A|\,|x_\nu-\zeta|^\alpha\quad(\nu\to\infty). \tag{5.10}$$

Так как $\alpha>0$, то

$$\frac{f(x_\nu)}{x_\nu-\zeta}\to0\quad(\nu\to\infty),$$

и потому

$$1-\left(1-\frac{f(x_\nu)}{x_\nu-\zeta}\right)^\alpha\sim\alpha\frac{f(x_\nu)}{x_\nu-\zeta}.$$

Теперь из (5.9) и (5.10) вытекает, что

$$1-\left(\frac{x_{\nu+1}-\zeta}{x_\nu-\zeta}\right)^\alpha \sim \alpha|A|\,|x_\nu-\zeta|^\alpha. \tag{5.11}$$

Полагая

$$q_\nu=|A|\,|x_\nu-\zeta|^\alpha,$$

перепишем (5.11) в виде

$$1-\frac{q_{\nu+1}}{q_\nu}\sim \alpha q_\nu.$$

Отсюда следует, что $q_{\nu+1}\sim q_\nu$,

$$\frac{1}{q_\nu}\left(1-\frac{q_{\nu+1}}{q_\nu}\right)\to\alpha \quad (\nu\to\infty)$$

и, наконец,

$$\frac{1}{q_{\nu+1}}-\frac{1}{q_\nu}=\frac{1}{q_{\nu+1}}\left(1-\frac{q_{\nu+1}}{q_\nu}\right)\to\alpha \quad (\nu\to\infty). \tag{5.12}$$

8. Чтобы закончить доказательство, воспользуемся следующей теоремой Коши:

если $d_\nu\to\alpha$ при $\nu\to\infty$, то последовательность

$$\frac{d_1}{1},\ \frac{d_1+d_2}{2},\ \ldots,\ \frac{d_1+d_2+\ldots+d_n}{n},\ \ldots$$

сходится к α.

Применяя теорему Коши к последовательности

$$d_\nu=\frac{1}{q_{\nu+1}}-\frac{1}{q_\nu}$$

и замечая, что

$$d_1+\ldots+d_n=$$

$$=\left(\frac{1}{q_2}-\frac{1}{q_1}\right)+\ldots+\left(\frac{1}{q_{n+1}}-\frac{1}{q_n}\right)=\frac{1}{q_{n+1}}-\frac{1}{q_1},$$

находим

$$\frac{\frac{1}{q_{\nu+1}}-\frac{1}{q_1}}{\nu}\to\alpha;$$

$$\frac{1}{\nu q_{\nu+1}}\to\alpha.$$

Но $q_{\nu+1} \sim q_\nu$, и потому

$$\frac{1}{\nu q_\nu} \to \alpha,$$

откуда

$$\nu \,|\, x_\nu - \zeta \,|^\alpha \to \frac{1}{\alpha \,|\, A \,|},$$

что равносильно доказываемому соотношению (5.8).

9. Если $\alpha \geqslant 1$, т. е. если ζ является корнем не менее чем второй кратности, то показатель степени у множителя $\nu^{1/\alpha}$ в (5.8) не превосходит 1 и $|x_\nu - \zeta|$ сходится к 0 очень медленно. (Индекс эффективности равен 1 при любом $\alpha > 0$.) Таким образом, в этом случае пользоваться рассматриваемым методом для расчетов нецелесообразно. Эффективные методы расчета для случая кратных корней будут даны в гл. 8 и в приложении V.

Связь regula falsi с теорией итерации

10. Итерацию по regula falsi

$$x_{\nu+1} = \frac{af(x_\nu) - x_\nu f(a)}{f(x_\nu) - f(a)} \tag{5.13}$$

можно рассматривать как итерацию с итерирующей функцией

$$\psi(x) = \frac{af(x) - xf(a)}{f(x) - f(a)}. \tag{5.14}$$

По лемме 4.2 корень ζ будет точкой притяжения, если

$$|\psi'(\zeta)| = \left|1 - f'(\zeta)\frac{a-\zeta}{f(a)}\right| < 1, \tag{5.15}$$

т. е. если

$$0 < \frac{f'(\zeta)}{\dfrac{f(a) - f(\zeta)}{a - \zeta}} < 2. \tag{5.16}$$

Случай $|\psi'(\zeta)| = 1$ является исключительным и здесь не рассматривается. Условие (5.16) означает, что угловой коэффициент хорды, соединяющей точки $(a, f(a))$ и $(\zeta, 0)$,

должен иметь тот же знак, что и $f'(\zeta)$, и по абсолютной величине должен быть больше, чем $\frac{1}{2}|f'(\zeta)|$. Это положение мы иллюстрируем рисунком 3.

На рис. 3 точка $(a, f(a))$ должна быть выбрана так, чтобы хорда, соединяющая ее с точкой $(\zeta, 0)$, не попала

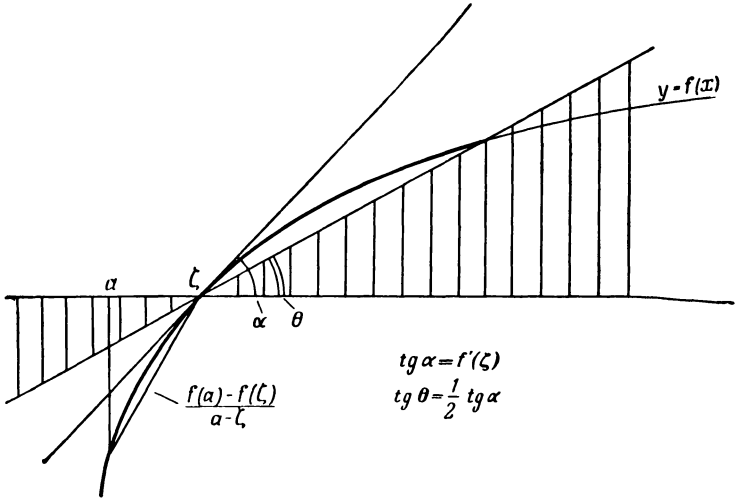

Рис. 3.

в заштрихованную область. Другими словами, a должно быть выбрано так, чтобы точка $(a, f(a))$ лежала на выделенной части кривой. Это дает правило, по которому можно выбрать точку a даже на большом расстоянии от ζ и все еще быть уверенным в сходимости метода. Если же наша хорда попадает внутрь заштрихованной области, то $|\psi'(\zeta)| > 1$ и ζ есть точка отталкивания.

ГЛАВА 6

МЕТОД НЬЮТОНА

Идея метода Ньютона

1. Пусть заданы *два совпадающих узла интерполяции* $x_1 = x_2 = a$ и значения $f(a)$, $f'(a)$. Как было сказано выше (см. стр. 12), линейный интерполяционный многочлен в этом случае является многочленом Тейлора

$$f(a) + (x - a) f'(a). \tag{6.1}$$

Приравнивая его нулю и разрешая относительно x, находим

$$x = a - \frac{f(a)}{f'(a)}. \tag{6.2}$$

Это выражение дает приближенное значение корня уравнения $f(x) = 0$, если корень лежит в окрестности точки a. Идея метода впервые была высказана Ньютоном, но формула (6.2) принадлежит Рафсону и в Англии известна как формула Ньютона — Рафсона [1]).

Если $a = x_0$ дает первое приближение к корню, то формула (6.2) дает второе приближение $x = x_1$. Вообще итерация

$$x_{\nu+1} = x_\nu - \frac{f(x_\nu)}{f'(x_\nu)} \tag{6.3}$$

дает последовательность приближений x_ν ($\nu = 1, 2, \ldots$). Мы рассмотрим следующие вопросы. Можно ли из свойств последовательности $\{x_\nu\}$ сделать вывод, что корень лежит вблизи наших приближений? Сходится ли последовательность к этому корню, и если да, то насколько хорошее приближение к корню дает x_ν?

---

[1]) См. прим. перев. на стр. 41.

Применение обратной интерполяции

2. Эти вопросы можно было бы исследовать с помощью многочлена (6.1), но быстрее ведет к результату рассмотрение обратной функции.

Пусть $f'(x) \neq 0$ в рассматриваемой окрестности точки x_0. Положим $y = f(x)$, $x = \varphi(y)$, $y_0 = f(x_0)$, $\varphi(y_0) = x_0$, $\varphi'(y_0) = \frac{1}{f'(x_0)}$. Тогда линейный интерполяционный многочлен для $\varphi(y)$ есть

$$L(y) = x_0 + (y - y_0)\frac{1}{f'(x_0)}. \tag{6.4}$$

По формуле (1.5) имеем

$$\varphi(y) - L(y) = \frac{\varphi''(\eta)}{2}(y - y_0)^2. \tag{6.5}$$

Если $\zeta = \varphi(0)$ — искомый корень, а x_0 — начальное приближение, то, подставляя в (6.5) $y = 0$, $L(0) = x_1$, получим погрешность приближения x_1:

$$\zeta - x_1 = \left[\frac{-f''(\xi)}{2f'(\xi)^3}\right] f(x_0)^2, \quad \xi \in (\zeta, x_0). \tag{6.6}$$

3. Заметим, что $\zeta - x_1$ квадратично зависит от значения $f(x_0)$. Для оценки погрешности (6.6) надо знать верхнюю грань $|f''(x)|$ и нижнюю грань $|f'(x)|$. Пусть в рассматриваемом интервале $|f'(x)| \geqslant m_1 > 0$, $|f''(x)| \leqslant M_2$; тогда

$$|\zeta - x_1| \leqslant \frac{M_2}{2m_1^3} f(x_0)^2 = k. \tag{6.7}$$

Значение $f(x_0)$ вычисляется при нахождении x_1 и непосредственно подставляется в эту оценку.

Если $f''(x)$ не меняет знака в исходной окрестности точки ζ, то знак выделенного в (6.6) множителя также не меняется, и мы можем получить новую окрестность, в которой лежит корень ζ. Если этот выделенный множитель положителен, то ζ лежит на сегменте $[x_1, x_1 + k]$, если же он отрицателен, то ζ лежит на сегменте $[x_1 - k, x_1]$.

В проведенном исследовании мы предполагали, что корень ζ *существует* и *лежит в окрестности точки* x_0. На практике часто приступают к вычислениям, не зная, верно ли это предположение. Такое поведение оправдано, ибо

часто после первых шагов расчета можно выяснить, существует ли корень в рассматриваемой окрестности и сходится ли построенная нами последовательность. Более подробно мы рассмотрим этот вопрос в гл. 7.

Здесь же мы рассмотрим вопрос о *быстроте сходимости*. Подставляя в (6.6) вместо $f(x_0)$ выражение

$$f(x_0) = f(x_0) - f(\zeta) = (x_0 - \zeta) f'(\xi_0), \quad \xi_0 \in (x_0, \zeta),$$

мы получим

$$\zeta - x_1 = \left[-\frac{1}{2} \frac{f''(\xi)}{f'(\xi)^3} f'(\xi_0)^2 \right] (\zeta - x_0)^2; \quad \xi, \xi_0 \in (x_0, \zeta), \tag{6.8}$$

откуда

$$\frac{\zeta - x_1}{(\zeta - x_0)^2} \to -\frac{1}{2} \frac{f''(\zeta)}{f'(\zeta)} \quad (x_0 \to \zeta). \tag{6.9}$$

Таким образом, наша аппроксимация, вообще говоря, улучшается *квадратично* на каждом шаге. Если выделенный в (6.8) множитель по модулю не превосходит 1 и если приближение x_0 дает n верных десятичных знаков корня, то приближение x_1 дает по меньшей мере $2n$ верных десятичных знаков.

Сравнение метода секущих с методом Ньютона

4. В гл. 3, п. 11, мы показали, что индекс эффективности метода секущих равен 1,618....

При расчете по методу Ньютона каждый шаг требует вычисления f и f', т. е. каждый шаг осуществляется посредством двух горнеров; следовательно, индекс эффективности метода Ньютона есть $\sqrt{2} = 1{,}414\ldots$. В этом смысле проведение итерации по формуле (3.15) *лучше*, чем по методу Ньютона.

Это преимущество, однако, имеет место лишь при решении алгебраических и трансцендентных уравнений. Метод секущих не может быть просто обобщен на другие теории. С другой стороны, принцип метода Ньютона находит применение в теориях дифференциальных уравнений, интегральных уравнений, функциональных уравнений и во многих других отделах анализа.

Успех метода зависит, конечно, и от выбора x_0. В этом отношении известно немногое.

ГЛАВА 7

ФУНДАМЕНТАЛЬНЫЕ ТЕОРЕМЫ СУЩЕСТВОВАНИЯ ДЛЯ МЕТОДА НЬЮТОНА

Оценки ошибок a priori и a posteriori

1. Иногда бывает возможным получить информацию о корне и оценки ошибок, исходя из начальных данных; например, если $f(x)$ непрерывна на $[a, b]$ и меняет на этом сегменте свой знак, то корень лежит на $[a, b]$. Оценки такого рода называют оценками a priori. На практике мы часто начинаем вычисления, не будучи уверенными в существовании корня. Используя результаты наших вычислений, мы можем получить оценки, которые, как правило, намного лучше, чем оценки a priori. Такие оценки называются оценками a posteriori. Оценки a priori бывают завышенными в сотни и тысячи раз. Оценки a posteriori могут давать верный порядок оцениваемой величины.

Фундаментальные теоремы существования

2. Мы сформулируем сразу две теоремы существования: для действительной функции действительного переменного и для аналитической функции комплексного переменного. Доказательства, за исключением одного пункта (доказательство единственности), аналогичны. В комплексном случае сегмент $J_0 : [x_0, x_0 + 2h_0]$ мы заменяем на круг $K_0 : |z - z_1| \leqslant |h_0|$, где z, z_1 — комплексные числа. Мы подробно докажем теорему для действительной функции $f(x)$ действительного переменного (теорема 7.2) и не будем отмечать очевидных изменений, которые необходимы при доказательстве теоремы для случая комплексного переменного (теорема 7.1), за исключением доказательства единственности.

В следующих теоремах *мы не делаем никаких предположений относительно существования корня.*

3. Теорема 7.1. Пусть $f(z)$ — комплексная функция комплексного переменного z, определенная в некоторой окрестности точки z_0, и пусть

$$f(z_0)f'(z_0) \neq 0, \quad h_0 = -\frac{f(z_0)}{f'(z_0)}, \quad z_1 = z_0 + h_0.$$

Пусть $f(z)$ аналитична в круге $K_0: |z - z_1| \leqslant |h_0|$ и $\max\limits_{K_0} |f''(z)| = M$, причем $2|h_0|M \leqslant |f'(z_0)|$. Построим, начиная с z_0, последовательность z_ν по рекуррентной формуле

$$z_{\nu+1} = z_\nu - \frac{f(z_\nu)}{f'(z_\nu)} \quad (\nu = 0, 1, \ldots).$$

Тогда все z_ν лежат в K_0 и

$$z_\nu \to \zeta \quad (\nu \to \infty),$$

где ζ — единственный в K_0 корень уравнения $f(z) = 0$. Корень ζ является простым, если он не лежит на границе K_0. Кроме того, справедливы оценки

$$\frac{|z_{\nu+1} - z_\nu|}{|z_\nu - z_{\nu-1}|^2} \leqslant \frac{M}{2|f'(z_\nu)|} \quad (\nu = 1, 2, \ldots),$$

$$|\zeta - z_{\nu+1}| \leqslant \frac{M}{2|f'(z_\nu)|} |z_\nu - z_{\nu-1}|^2 \quad (\nu = 1, 2, \ldots).$$

4. Теорема 7.2. Пусть $f(x)$ — действительная функция действительного переменного x;

$$f(x_0)f'(x_0) \neq 0, \quad h_0 = -\frac{f(x_0)}{f'(x_0)}, \quad x_1 = x_0 + h_0.$$

Пусть на сегменте $J_0: [x_0, x_0 + 2h_0]$ существует $f''(x)$ и $\max\limits_{|J_0|} |f''(x)| = M$, причем

$$2|h_0|M \leqslant |f'(x_0)|. \tag{7.1}$$

Построим, начиная с x_0, последовательность x_ν по рекуррентной формуле

$$x_{\nu+1} = x_\nu - \frac{f(x_\nu)}{f'(x_\nu)} \quad (\nu = 0, 1, \ldots).$$

Тогда все x_ν лежат в J_0 и

$$x_\nu \to \zeta \quad (\nu \to \infty), \tag{7.2}$$

где ζ — единственный в J_0 корень уравнения $f(x) = 0$.

Корень ζ является простым, если он не лежит на границе J_0. Кроме того, справедливы оценки

$$\text{a)}\quad \frac{|x_{\nu+1} - x_\nu|}{|x_\nu - x_{\nu-1}|^2} \leqslant \frac{M}{2|f'(x_\nu)|} \quad (\nu = 1, 2, \ldots),$$

$$\text{b)}\quad |\zeta - x_{\nu+1}| \leqslant \frac{M}{2|f'(x_\nu)|} |x_\nu - x_{\nu-1}|^2 \quad (\nu = 1, 2, \ldots).$$

5. Доказательство. Из формул

$$f'(x) - f'(x_0) = \int_{x_0}^{x} f''(x)\,dx,$$

$$|f'(x) - f'(x_0)| \leqslant |x - x_0| M \tag{7.3}$$

и условия (7.1) получаем оценку

$$|f'(x_1) - f'(x_0)| \leqslant |h_0| M \leqslant \frac{|f'(x_0)|}{2}; \tag{7.4}$$

из этой оценки следует

$$|f'(x_1)| \geqslant |f'(x_0)| - |f'(x_1) - f'(x_0)| \geqslant \frac{|f'(x_0)|}{2}. \tag{7.5}$$

Далее, интегрируя по частям, получаем

$$\int_{x_0}^{x_1} (x_1 - x) f''(x)\,dx = -(x_1 - x_0) f'(x_0) + f(x_1) - f(x_0) =$$

$$= -h_0 f'(x_0) - f(x_0) + f'(x_1) = f(x_1),$$

$$f(x_1) = \int_{x_0}^{x_1} (x_1 - x) f''(x)\,dx. \tag{7.6}$$

что дает возможность оценить $f(x_1)$ [1]).

---

[1]) Выражение (7.6) есть представление $f(x_1)$ по формуле Тейлора с остаточным членом в интегральной форме:

$$f(x_1) = f(x_0) + h_0 f'(x_0) + \int_{x_0}^{x_1} (x_1 - x) f''(x)\,dx.$$

Оценку (7.7) для случая действительного переменного можно получить и непосредственно из формулы Тейлора вида

$$f(x_1) = f(x_0) + h_0 f'(x_0) + \frac{1}{2} h_0^2 f''(x_0 + \vartheta h_0) \quad (0 < \vartheta < 1).$$ — *Прим. перев.*

Введем новую переменную t, положив

$$x = x_0 + th_0, \quad x_1 - x = h_0(1-t), \quad dx = h_0\,dt:$$

$$f(x_1) = h_0^2 \int_0^1 (1-t) f''(x_0 + th_0)\,dt.$$

6. Учитывая, что $1 - t \geqslant 0$, получаем

$$|f(x_1)| \leqslant |h_0|^2 \cdot M \int_0^1 (1-t)\,dt = \frac{|h_0|^2 M}{2}. \tag{7.7}$$

Положим $h_1 = -\dfrac{f(x_1)}{f'(x_1)}$, $x_2 = x_1 + h_1$. Применяя (7.5) и (7.7), получим

$$|h_1| \leqslant \frac{|h_0|^2 M}{|f'(x_0)|}, \tag{7.8}$$

$$\frac{2M|h_1|}{|f'(x_1)|} \leqslant \frac{2|h_0|^2 M^2}{|f'(x_0)|\,|f'(x_1)|} \leqslant \frac{2|h_0|^2 M^2}{|f'(x_0)|\,\frac{1}{2}|f'(x_0)|} =$$

$$= \left(\frac{2|h_0|M}{|f'(x_0)|}\right)^2.$$

В силу условия (7.1) правая часть этого неравенства не превосходит единицы.

Следовательно,

$$2|h_1|M \leqslant |f'(x_1)|. \tag{7.9}$$

С другой стороны, по (7.8)

$$|h_1| \leqslant \frac{1}{2}|h_0|\left(\frac{2|h_0|M}{|f'(x_0)|}\right) \leqslant \frac{1}{2}|h_0|. \tag{7.10}$$

Из (7.10) видим, что x_2 удалено от x_1 на расстояние, не превосходящее $\frac{1}{2}|h_0|$, и потому x_2 остается на сегменте J_0. Далее, сегмент $J_1 : [x_1, x_1 + 2h_1]$ принадлежит J_0, $J_1 \subset J_0$.

7. Положим вообще

$$x_{\nu+1} = x_\nu + h_\nu, \quad h_\nu = -\frac{f(x_\nu)}{f'(x_\nu)}. \tag{7.11}$$

Соотношение (7.9) показывает, что условия нашей теоремы будут выполнены, если мы заменим x_0 и h_0 на x_1 и h_1 соответственно; следовательно, эти условия остаются справедливыми для x_ν и h_ν $(\nu = 0, 1, \ldots)$.

Обозначим сегмент $[x_\nu, x_\nu + 2h]$ через J $(\nu = 0, 1, \ldots)$. Мы имеем последовательность вложенных сегментов $J_{\nu+1} \subset \subset J_\nu$, причем длина сегмента $J_{\nu+1}$ не превосходит половины длины J_ν. Известно, что такая последовательность стягивается к точке ζ, лежащей в J_0:

$$x_\nu \to \zeta, \quad \zeta \in J_0. \tag{7.12}$$

Чтобы доказать, что ζ является корнем $f(x)$, перейдем в (7.11) к пределу при $x_\nu \to \zeta$:

$$h_\nu = x_{\nu+1} - x_\nu \to 0; \quad f(x_\nu) = -h_\nu f'(x_\nu) \to 0,$$

откуда

$$f(\zeta) = 0. \tag{7.13}$$

8. Из (7.3) для всех x, лежащих *внутри* (J_0), т. е. таких, что $(|x - x_0| < 2|h_0|)$, имеем

$$|f'(x) - f'(x_0)| \leqslant |x - x_0| M < 2|h_0| M \leqslant |f'(x_0)|. \tag{7.14}$$

Отсюда видно, что $f'(x) \neq 0$, если x лежит внутри (J_0). Поэтому ζ является простым корнем, если $\zeta \in (J_0)$.

Мы докажем теперь, что ζ — единственный корень на J_0. *В случае действительного переменного* $f'(x)$ не обращается в нуль на (J_0). Поэтому $f(x)$ строго монотонна на J_0 и имеет там только один корень.

9. *В случае комплексного переменного* требуется, очевидно, другое доказательство. Пусть ζ^* — корень $f(z)$, лежащий *внутри* K_0. Для доказательства единственности корня достаточно показать, что $\zeta^* = \lim\limits_{\nu \to \infty} z_\nu$, т. е. что

$$|\zeta^* - z_\nu| \to 0 \quad (\nu \to \infty).$$

Интегрируя по частям, получаем

$$\int_{z_\nu}^{\zeta^*} (z-\zeta^*) f''(z)\,dz = (z-\zeta^*) f'(z) \Big|_{z_\nu}^{\zeta^*} - \int_{z_\nu}^{\zeta^*} f'(z)\,dz =$$

$$= f'(z_\nu)\left[\zeta^* - z_\nu + \frac{f(z_\nu)}{f'(z_\nu)}\right] = f'(z_\nu)(\zeta^* - z_{\nu+1}),$$

$$f'(z_\nu)(\zeta^* - z_{\nu+1}) = \int_{z_\nu}^{\zeta^*} (z-\zeta^*) f''(z)\,dz.$$

В этом интеграле заменим переменную:

$$z = z_\nu + t(\zeta^* - z_\nu); \quad z - \zeta^* = (z_\nu - \zeta^*)(1-t);$$

$$dz = (\zeta^* - z_\nu)\,dt;$$

$$f'(z_\nu)(\zeta^* - z_{\nu+1}) =$$

$$= -(\zeta^* - z_\nu)^2 \int_0^1 (1-t) f''[z_\nu + t(\zeta^* - z_\nu)]\,dt.$$

Отсюда получаем оценку

$$|f'(z_\nu)|\,|\zeta^* - z_{\nu+1}| \leqslant |\zeta^* - z_\nu|^2 \cdot \frac{M}{2}.$$

Полагая $d_\nu = |\zeta^* - z_\nu|$, запишем последнее неравенство в виде

$$\frac{d_{\nu+1}}{d_\nu^2} = \frac{|\zeta^* - z_{\nu+1}|}{|\zeta^* - z_\nu|^2} \leqslant \frac{M}{2|f'(z_\nu)|}. \tag{7.15}$$

10. Так как $\zeta^* \in K_0$, то $d_0 = |\zeta^* - z_0| \leqslant 2|h_0|$, и в силу условия (7.1)

$$\frac{d_0 M}{|f'(z_0)|} \leqslant 1.$$

Поэтому из (7.15) при $\nu = 0$ получаем

$$2\frac{d_1}{d_0} \leqslant d_0 \frac{M}{|f'(z_0)|} \leqslant 1, \tag{7.16}$$

Далее, условия теоремы выполнены в каждой точке z_ν ($\nu = 0, 1, \ldots$), поэтому в каждой точке z_ν выполнено неравенство типа (7.5):

$$|f'(z_{\nu+1})| \geqslant \frac{1}{2}|f'(z_\nu)|.$$

Повторно применяя это неравенство, получаем

$$|f'(z_\nu)| \geqslant \frac{1}{2^\nu}|f'(z_0)|,$$

и, следовательно, по (7.15),

$$\frac{d_{\nu+1}}{d_\nu^2} \leqslant \frac{M}{2|f'(z_\nu)|} \leqslant 2^{\nu-1}\frac{d_0 M}{|f'(z_0)|}\frac{1}{d_0} \leqslant 2^{\nu-1}\frac{1}{d_0},$$

$$2^{\nu+1}\frac{d_{\nu+1}}{d_0} \leqslant 2^{2\nu}\frac{d_\nu^2}{d_0^2} = \left(2^\nu\frac{d_\nu}{d_0}\right)^2. \tag{7.17}$$

Так как $2\frac{d_1}{d_0} \leqslant 1$, то из (7.17) по индукции находим

$$2^\nu\frac{d_\nu}{d_0} \leqslant 1.$$

Поэтому

$$|\zeta^* - z_\nu| \leqslant \frac{1}{2^\nu}|\zeta^* - z_0|,$$

и

$$z_\nu \to \zeta^* \quad (\nu \to \infty).$$

11. Докажем теперь неравенства (a) и (b) теоремы 7.2. Условия теоремы выполнены для всех x_ν ($\nu = 0, 1, \ldots$), поэтому для них справедливо неравенство типа (7.7):

$$|f(x_\nu)| \leqslant \frac{M}{2}|h_{\nu-1}|^2,$$

откуда в силу (7.11)

$$|h_\nu| = \frac{|f(x_\nu)|}{|f'(x_\nu)|} \leqslant \frac{M|h_{\nu-1}|^2}{2|f'(x_\nu)|}. \tag{7.18}$$

Это последнее неравенство равносильно (a).

Наконец, заметим, что ζ лежит на сегменте с центром $x_{\nu+1}$ и радиусом $|h_\nu|$, т. е. $|\zeta - x_{\nu+1}| \leqslant |h_\nu|$. Применяя здесь неравенство (7.18), получим (b). Теорема доказана полностью.

Если ζ лежит на границе, то $\zeta = x_0 + 2h_0$; в этом случае легко показать, что $f(x)$ является квадратичным многочленом с двойным корнем ζ, т. е. что $f(x) = a(x - \zeta)^2$.

Теперь ясно, что, проводя расчет по методу Ньютона, мы можем быть уверены в существовании корня, если на одном из шагов расчета выполнено условие $2|h_\nu| M \leqslant \leqslant |f'(x_\nu)|$. Существенно, что $f'(x_0) \neq 0$. Если ζ — простой корень и x_0 выбрано достаточно близко к ζ, то условия нашей теоремы будут выполнены.

Если ζ — кратный корень и x_0 выбрано достаточно близко к ζ, то $f'(x_0)$ будет очень мало. При этом последовательность, построенная по методу Ньютона, будет сходиться очень медленно (как в случае, рассмотренном в гл. 5) или вообще не будет сходиться. Модификации метода для этого случая рассматриваются в следующей главе.

Полное исследование сходимости в случае квадратичного многочлена дано в приложении VI.

Некоторые модификации и улучшения метода Ньютона даны в приложениях VII и IX, п. 6.

ГЛАВА 8

АНАЛОГ МЕТОДА НЬЮТОНА ДЛЯ КРАТНЫХ КОРНЕЙ

1. Формула, аналогичная формуле Ньютона и применимая в случае кратных корней, дана Шредером [5][1]).

Допустим, что корень имеет *точную кратность* p. В этом случае мы заменим формулу Ньютона на

$$x_{\nu+1} = x_\nu - p\frac{f(x_\nu)}{f'(x_\nu)} \quad (\nu = 1, 2, \ldots)^{2)}. \tag{8.1}$$

При такой модификации трудности, связанные с кратностью корня, устраняются и сходимость остается сверхлинейной.

Теорема 8.1. Пусть $f(x)$ имеет корень ζ точной кратности p; пусть $f^{(p+1)}(x)$ существует и непрерывна в некоторой окрестности точки ζ. Если точка x_1 достаточно близка к ζ, то последовательность $\{x_\nu\}$, построенная по формуле (8.1), сходится к ζ и

$$\zeta - x_{\nu+1} \sim -(\zeta - x_\nu)^2 \frac{f^{(p+1)}(\zeta)}{p(p+1)f^{(p)}(\zeta)} \quad (\nu \to \infty). \tag{8.2}$$

2. Заметим, что при $p = 1$ формула (8.2) дает уже известный результат (6.9) для метода Ньютона.

Доказательство теоремы 8.1. По формуле (8.1)

$$\zeta - x_{\nu+1} = \zeta - x_\nu + p\frac{f(x_\nu)}{f'(x_\nu)} = \frac{F(x_\nu)}{f'(x_\nu)}, \tag{8.3}$$

где

$$F(x) = pf(x) - f'(x)(x - \zeta). \tag{8.4}$$

---

[1]) Номера в квадратных скобках относятся к библиографии в конце книги.

[2]) Формула (8.1) получается из формулы (6.3) заменой $f(x)$ на $|f(x)|^{1/p}$. [Той же заменой можно получить оценку (8.2) из оценки (6.9).—*Прим. перев.*]

При доказательстве мы используем разложение Тейлора

$$\varphi(x) = \sum_{\nu=0}^{n} \frac{(x-\zeta)^\nu}{\nu!} \varphi^{(\nu)}(\zeta) + R_{n+1}, \tag{8.5}$$

где

$$R_{n+1} = \frac{(x-\zeta)^{n+1}}{(n+1)!} \varphi^{(n+1)}(\xi), \quad \xi \in (\zeta, x), \tag{8.6}$$

или

$$R_{n+1} = \int_\zeta^x \frac{(x-t)^n}{n!} \varphi^{(n+1)}(t)\, dt. \tag{8.7}$$

Вычислим производные от $F(x)$ с помощью формулы Лейбница дифференцирования произведения:

$$\begin{aligned} F^{(\nu)}(x) &= pf^{(\nu)}(x) - [f^{(\nu+1)}(x)(x-\zeta) + \nu f^{(\nu)}(x)] = \\ &= (p-\nu) f^{(\nu)}(x) - (x-\zeta) f^{(\nu+1)}(x). \end{aligned} \tag{8.8}$$

Так как корень ζ, имеет кратность p

$$f^{(\nu)}(\zeta) = 0 \quad (\nu = 0, 1, \ldots, p-1); \quad f^{(p)}(\zeta) \neq 0,$$

и поэтому

$$F^{(\nu)}(\zeta) = 0 \quad (\nu = 0, 1, \ldots, p). \tag{8.9}$$

3. Применим разложение (8.5) к $f'(x)$:

$$f'(x) = \sum_{\nu=0}^{p-2} \frac{(x-\zeta)^\nu}{\nu!} f^{(\nu+1)}(\zeta) + R_{p-1} = R_{p-1},$$

и по (8.6)

$$f'(x_\nu) = \frac{(x_\nu - \zeta)^{p-1}}{(p-1)!} f^{(p)}(\xi_1), \quad \xi_1 \in (\zeta, x_\nu). \tag{8.10}$$

Далее, применим разложение (8.5) к $F(x)$:

$$F(x) = \sum_{\nu=0}^{p-1} \frac{(x-\zeta)^\nu}{\nu!} F^{(\nu)}(\zeta) + R_p.$$

Поэтому в силу (8.7), (8.8) и (8.9)

$$\begin{aligned} F(x) &= \int_\zeta^x \frac{(x-t)^{p-1}}{(p-1)!} F^{(p)}(t)\, dt = \\ &= -\frac{1}{(p-1)!} \int_\zeta^x [(x-t)^{p-1}(t-\zeta)] f^{(p+1)}(t)\, dt. \end{aligned}$$

4. Множитель, заключенный в квадратные скобки, не меняет знака на промежутке интегрирования, что дает возможность применить обобщенную теорему о среднем:

$$F(x) = -\frac{f^{(p+1)}(\xi_2)}{(p-1)!}\int_{\zeta}^{x}(x-t)^{p-1}(t-\zeta)\,dt, \quad \xi_2 \in (\zeta, x).$$

Интегрируя по частям, получаем

$$\int_{\zeta}^{x}(x-t)^{p-1}(t-\zeta)\,dt = -\frac{(x-t)^p(t-\zeta)}{p}\Bigg|_{\zeta}^{x} + \int_{\zeta}^{x}\frac{(x-t)^p}{p}\,dt =$$

$$= \frac{(x-\zeta)^{p+1}}{p(p+1)}.$$

Поэтому

$$F(x_\nu) = -\frac{f^{(p+1)}(\xi_2)}{(p-1)!}\,\frac{(x_\nu-\zeta)^{p+1}}{p(p+1)}, \quad \xi_2 \in (\zeta, x_\nu).$$

Подставляя это выражение вместе с (8.10) в (8.3), находим

$$\zeta - x_{\nu+1} = -\frac{(\zeta-x_\nu)^2 f^{(p+1)}(\xi_2)}{p(p+1)f^{(p)}(\xi_1)}, \quad \xi_1, \xi_2 \in (\zeta, x_\nu). \quad (8.11)$$

Если теперь $x_\nu \to \zeta$, то $\xi_1 \to \zeta$, $\xi_2 \to \zeta$, и отсюда непосредственно следует (8.2)[1]).

---

[1]) Этот результат можно получить из (8.3) более экономно, применяя разложение Тейлора (8.5), (8.6) к $f(x_\nu)$ и $f'(x_\nu)$:

$$F(x_\nu) = pf(x_\nu) - f'(x_\nu)(x_\nu-\zeta) =$$

$$= p\left[\frac{f^{(p)}(\zeta)}{p!}(x_\nu-\zeta)^p + \frac{f^{(p+1)}(\xi_3)}{(p+1)!}(x_\nu-\zeta)^{p+1}\right] -$$

$$-(x_\nu-\zeta)\left[\frac{f^{(p)}(\zeta)}{(p-1)!}(x_\nu-\zeta)^{p-1} + \frac{f^{(p+1)}(\xi_4)}{p!}(x_\nu-\zeta)^p\right] =$$

$$= \frac{(x_\nu-\zeta)^{p+1}}{(p-1)!}\left[\frac{1}{p+1}f^{(p+1)}(\xi_3) - \frac{1}{p}f^{(p+1)}(\xi_4)\right] \sim$$

$$\sim -\frac{(x_\nu-\zeta)^{p+1}}{(p-1)!}\,\frac{f^{(p+1)}(\zeta)}{p(p+1)} \text{ при } x_\nu \to \zeta \quad (\xi_3, \xi_4 \to \zeta);$$

$$f'(x_\nu) \sim \frac{(x_\nu-\zeta)^{p-1}}{(p-1)!}f^{(p)}(\zeta). \text{ — } \textit{Прим. перев.}$$

Положим теперь

$$\inf_{(\zeta,\, x_\nu)} |f^{(p)}(x)| = m_p, \quad \sup_{(\zeta,\, x_\nu)} |f^{(p+1)}(x)| = M_{p+1} \ ;$$

тогда из (8.11) получим

$$|\zeta - x_{\nu+1}| \leqslant |\zeta - x_\nu|^2 \frac{M_{p+1}}{p(p+1)m_p}. \tag{8.12}$$

Замечание. Часто бывает трудно определить m_p. Хорошую оценку этой величины можно получить, зная M_{p+1} и производную $f^{(p)}(x_0)$ в некоторой точке x_0. Оценка m_p находится из неравенства

$$|f^{(p)}(x)| \geqslant |f^{(p)}(x_0)| - M_{p+1}|x - x_0|. \tag{8.13}$$

ГЛАВА 9

ГРАНИЦЫ ФУРЬЕ ПРИ ИТЕРАЦИЯХ НЬЮТОНА

1. В этой главе мы будем предполагать, что выполнены так называемые «условия Фурье»:

Функция $f(x)$ определена и имеет две производные на $J_0: x_0 \leqslant x \leqslant y_0$; ее корень ζ лежит на J_0, $f'(x) f''(x) \neq 0$ при $x \in J_0$, $f(x_0) f(y_0) < 0$ и $f(x_0) f''(x_0) > 0$.

Без ограничения общности можно считать, что

$$f(x_0) < 0 < f(y_0); \quad f'(x) > 0, \quad f''(x) < 0, \quad (x \in J_0) \tag{F}$$

(см. рис. 4).

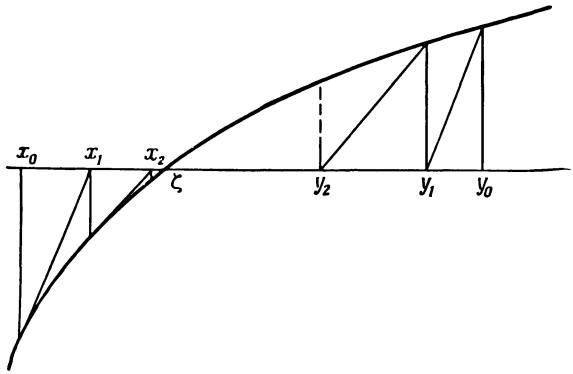

Рис. 4.

Пусть последовательность x_ν $(\nu = 0, 1, \ldots)$ определена по формуле Ньютона (6.3). Определим последовательность y_ν $(\nu = 0, 1, \ldots)$ по формуле

$$y_{\nu+1} = y_\nu - \frac{f(y_\nu)}{f'(x_\nu)}. \tag{9.1}$$

Заметим, что точка $y_{\nu+1}$ есть точка пересечения прямой, проходящей через точку $(y_\nu, f(y_\nu))$ с угловым коэффициентом $f'(x_\nu)$, и оси абсцисс.

Из рис. 4 видно, что последовательность $\{x_\nu\}$ монотонно стремится к числу $\xi \leqslant \zeta$, а $\{y_\nu\}$ — к числу $\eta \geqslant \zeta$. Переходя в формулах (6.3) и (9.1) к пределу при $\nu \to \infty$, получим

$$-\frac{f(\xi)}{f'(\xi)} = 0; \qquad -\frac{f(\eta)}{f'(\xi)} = 0.$$

Так как ζ — единственный корень $f(x)$ на J_0, то

$$\xi = \eta = \zeta,$$

т. е.

$$x_\nu \uparrow \zeta, \quad y_\nu \downarrow \zeta$$

и

$$(y_\nu - x_\nu) \downarrow 0 \qquad (\nu \to \infty). \tag{9.2}$$

Докажем теперь, что если $f''(x)$ непрерывна в (J_0), то

$$\frac{y_{\nu+1} - x_{\nu+1}}{(y_\nu - x_\nu)^2} \to -\frac{1}{2}\,\frac{f''(\zeta)}{f'(\zeta)} \qquad (\nu \to \infty), \tag{9.3}$$

т. е. расстояние между x_ν и y_ν убывает по квадратичному закону.

2. Положим

$$\lambda_\nu = \frac{y_\nu - \zeta}{\zeta - x_\nu}. \tag{9.4}$$

Если бы $\lambda_\nu \to 1$ при $\nu \to \infty$, то y_ν давали бы достаточно хорошее приближение к корню ζ. *Мы докажем, однако, что* $\lambda_\nu \to \infty$ *и даже что* $\frac{\lambda_{\nu+1}}{\lambda_\nu^2} \to 1$.

По формуле (9.1)

$$\frac{y_{\nu+1} - \zeta}{y_\nu - \zeta} = 1 - \frac{f(y_\nu)}{f'(x_\nu)(y_\nu - \zeta)} =$$

$$= \frac{1}{f'(x_\nu)}\left[f'(x_\nu) - \frac{f(\zeta) - f(y_\nu)}{\zeta - y_\nu}\right]. \tag{9.5}$$

В дальнейшем мы неоднократно будем представлять приращение функции в интегральной форме с помощью формулы

$$\frac{f(a)-f(b)}{a-b}=\int_0^1 f'[b+t(a-b)]\,dt=$$

$$=\int_0^1 f'[a+t(b-a)]\,dt \qquad (a\neq b). \tag{9.6}$$

В частности,

$$\frac{f(\zeta)-f(y_\nu)}{\zeta-y_\nu}=\int_0^1 f'[y_\nu+t(\zeta-y_\nu)]\,dt. \tag{9.7}$$

3. Учитывая, что

$$f'(x_\nu)=\int_0^1 f'(x_\nu)\,dt,$$

запишем (9.5) в интегральной форме:

$$\frac{y_{\nu+1}-\zeta}{y_\nu-\zeta}=-\frac{1}{f'(x_\nu)}\int_0^1\{f'[y_\nu+t(\zeta-y_\nu)]-f'(x_\nu)\}\,dt. \tag{9.8}$$

Применяя (9.6) к подинтегральной функции, получаем

$$f'[y_\nu+t(\zeta-y_\nu)]-f'(x_\nu)=$$

$$=[y_\nu-x_\nu+t(\zeta-y_\nu)]\int_0^1 f''\{x_\nu+u[y_\nu-x_\nu+t(\zeta-y_\nu)]\}\,du \tag{9.9}$$

Подставим это выражение в (9.8) и разделим на $(y_\nu-\zeta)$:

$$\frac{y_{\nu+1}-\zeta}{(y_\nu-\zeta)^2}=-\frac{1}{f'(x_\nu)}\int_0^1\left(\frac{y_\nu-x_\nu}{y_\nu-\zeta}-t\right)dt\int_0^1 f''\,du, \tag{9.10}$$

где $f''=f''\{x_\nu+u[y_\nu-x_\nu+t(\zeta-y_\nu)]\}$.

4. Так как множитель

$$\frac{y_\nu - x_\nu}{y_\nu - \zeta} - t = 1 + \frac{1}{\lambda_\nu} - t$$

положителен при $0 \leqslant t \leqslant 1$, то к интегралу (9.10) можно применить обобщенную теорему о среднем значении:

$$\frac{y_{\nu+1} - \zeta}{(y_\nu - \zeta)^2} = -\frac{f''(\xi)}{f'(x_\nu)} \int_0^1 \int_0^1 \left(1 + \frac{1}{\lambda_\nu} - t\right) du\, dt, \quad \xi \in (x_\nu, y_\nu),$$

откуда

$$\frac{y_{\nu+1} - \zeta}{(y_\nu - \zeta)^2} = -\frac{f''(\xi)}{f'(x_\nu)} \left(\frac{1}{2} + \frac{1}{\lambda_\nu}\right). \tag{9.11}$$

Из (9.11) при $\nu \to \infty$ получаем

$$\frac{y_{\nu+1} - \zeta}{(y_\nu - \zeta)^2} \cdot \frac{\lambda_\nu}{2 + \lambda_\nu} \to -\frac{1}{2} \frac{f''(\zeta)}{f'(\zeta)}. \tag{9.12}$$

С другой стороны, из (6.9)

$$\frac{\zeta - x_{\nu+1}}{(\zeta - x_\nu)^2} \to -\frac{1}{2} \frac{f''(\zeta)}{f'(\zeta)} \quad (\nu \to \infty). \tag{9.13}$$

Разделив (9.12) на (9.13), получаем

$$\frac{\lambda_{\nu+1}}{\lambda_\nu^2} \cdot \frac{\lambda_\nu}{2 + \lambda_\nu} \to 1, \text{ или } \frac{\lambda_{\nu+1}}{\lambda_\nu} \frac{1}{2 + \lambda_\nu} \to 1 \quad (\nu \to \infty). \tag{9.14}$$

Так как $\lambda_\nu > 0$, то из (9.14) следует, что при достаточно больших ν должно выполняться неравенство

$$\frac{\lambda_{\nu+1}}{\lambda_\nu} \frac{1}{2 + \lambda_\nu} \geqslant \frac{3}{4}; \quad \frac{\lambda_{\nu+1}}{\lambda_\nu} \geqslant \frac{3}{4}(2 + \lambda_\nu) \geqslant \frac{3}{2},$$

поэтому

$$\lambda_\nu \to \infty \quad (\nu \to \infty). \tag{9.15}$$

Теперь из (9.14) получаем

$$\frac{\lambda_{\nu+1}}{\lambda_\nu^2} \to 1 \quad (\nu \to \infty). \tag{9.16}$$

Наконец, из равенства

$$y_\nu - x_\nu = (\zeta - x_\nu)(1 + \lambda_\nu)$$

следует, что

$$\frac{y_{\nu+1} - x_{\nu+1}}{(y_\nu - x_\nu)^2} = \frac{\zeta - x_{\nu+1}}{(\zeta - x_\nu)^2} \frac{1 + \lambda_{\nu+1}}{(1 + \lambda_\nu)^2};$$

применяя здесь (9.13), (9.15) и (9.16), приходим к соотношению

$$\frac{y_{\nu+1} - x_{\nu+1}}{(y_\nu - x_\nu)^2} \to -\frac{1}{2} \frac{f''(\zeta)}{f'(\zeta)} \quad (\nu \to \infty), \text{ ч. т. д.}$$

ГЛАВА 10

ГРАНИЦЫ ДАНДЕЛЕНА ПРИ ИТЕРАЦИЯХ НЬЮТОНА

1. Предположим, как и в гл. 9, что на сегменте J_0, $x_0 \leqslant x \leqslant z_0$ выполнены условия Фурье и, в частности, условие (F) гл. 9 (с заменой y_0 на z_0). Построим последовательность x_ν ($\nu = 0, 1, \ldots$) по формуле Ньютона (6.3) и последовательность z_ν по формуле

$$z_{\nu+1} = z_\nu - f(z_\nu)\frac{z_\nu - x_\nu}{f(z_\nu) - f(x_\nu)} \qquad (\nu = 0, 1, \ldots). \qquad (10.1)$$

Формула (10.1) есть формула regula falsi для начальных точек x_ν и z_ν [1].

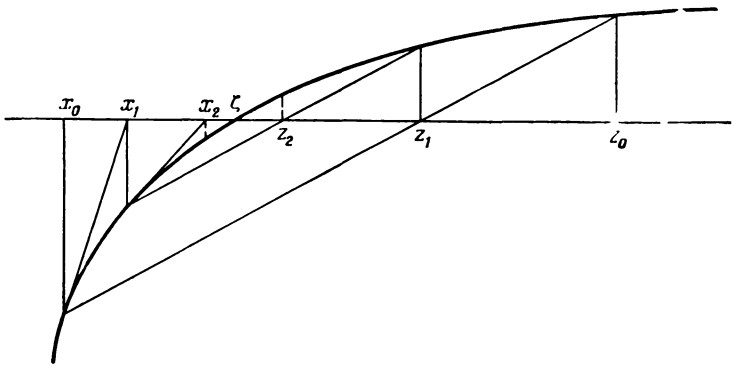

Рис. 5.

Из рис. 5 видно, что z_ν, монотонно убывая, стремятся к некоторому пределу η. Этот предел совпадает с кор-

[1]) Итерация по формулам (6.3) и (10.1) представляет собой комбинацию метода Ньютона с regula falsi и в русской литературе известна под названием «метод хорд и касательных».—*Прим. перев.*

нем ζ. Действительно, из формулы (10.1) следует, что при $\nu \to \infty$

$$f(z_\nu)\frac{z_\nu - x_\nu}{f(z_\nu) - f(x_\nu)} \to 0.$$

Но второй множитель остается при этом больше некоторого положительного числа:

$$\frac{z_\nu - x_\nu}{f(z_\nu) - f(x_\nu)} \geqslant \frac{1}{\min\limits_{x \in J_0} f'(x)} > 0;$$

поэтому $f(z_\nu) \to 0$, $f(\eta) = 0$, и так как в J_0 есть только один корень ζ, то

$$\eta = \zeta, \quad z_\nu \downarrow \zeta.$$

Теорема 10.1. Если $f''(x)$ непрерывна на J_0, то последовательность

$$\mu_\nu = \frac{z_\nu - \zeta}{\zeta - x_\nu} \quad (\nu = 0, 1, \ldots) \tag{10.2}$$

сходится при $\nu \to \infty$ к некоторому пределу $\Lambda(z_0)$. Эта сходимость равномерна по z_0 в $\zeta < z_0 \leqslant y_0$ при фиксированном y_0; предельная функция $\Lambda(z_0)$ непрерывна и, монотонно убывая, стремится к 0 при $z_0 \to \zeta$.

2. Доказательство. Мы имеем тождественно

$$\frac{z_{\nu+1} - \zeta}{z_\nu - \zeta} = 1 - \frac{f(z_\nu) - f(\zeta)}{z_\nu - \zeta}\,\frac{z_\nu - x_\nu}{f(z_\nu) - f(x_\nu)} = \frac{N}{D}, \tag{10.3}$$

где

$$N = \frac{f(x_\nu) - f(z_\nu)}{x_\nu - z_\nu} - \frac{f(\zeta) - f(z_\nu)}{\zeta - z_\nu}; \quad D = \frac{f(z_\nu) - f(x_\nu)}{z_\nu - x_\nu}. \tag{10.4}$$

Применяя преобразование (9.6), мы последовательно находим

$$\begin{gathered} D = \int_0^1 f'[x_\nu + t(z_\nu - x_\nu)]\,dt; \\ N = \int_0^1 \{f'[z_\nu + t(x_\nu - z_\nu)] - f'[z_\nu + t(\zeta - z_\nu)]\}\,dt \end{gathered} \tag{10.5}$$

и, наконец,

$$N = -(\zeta - x_\nu)\int_0^1\int_0^1 f''[z_\nu + t(x_\nu - z_\nu) + \\ + ut(\zeta - x_\nu)]\, t\, du\, dt. \tag{10.6}$$

3. С другой стороны,

$$\frac{\zeta - x_{\nu+1}}{\zeta - x_\nu} = 1 + \frac{f(x_\nu)}{f'(x_\nu)(\zeta - x_\nu)} = \\ = \frac{1}{f'(x_\nu)}\left[f'(x_\nu) - \frac{f(\zeta) - f(x_\nu)}{\zeta - x_\nu}\right].$$

Снова применяя два раза преобразование (9.6), находим

$$\frac{\zeta - x_{\nu+1}}{\zeta - x_\nu} = -\frac{1}{f'(x_\nu)}\int_0^1 \{f'[x_\nu + t(\zeta - x_\nu)] - f'(x_\nu)\}\, dt = \\ = -\frac{\zeta - x_\nu}{f'(x_\nu)}\int_0^1\int_0^1 f''[x_\nu + ut(\zeta - x_\nu)]\, t\, du\, dt.$$

Обозначим

$$S = \int_0^1\int_0^1 f''[x_\nu + ut(\zeta - x_\nu)]\, t\, du\, dt. \tag{10.7}$$

Тогда

$$\frac{\zeta - x_{\nu+1}}{\zeta - x_\nu} = -\frac{\zeta - x_\nu}{f'(x_\nu)} S. \tag{10.8}$$

Разделив (10.3) на (10.8), мы получим

$$\frac{\mu_{\nu+1}}{\mu_\nu} = \frac{\dfrac{(z_{\nu+1} - \zeta)}{(\zeta - x_{\nu+1})}}{\dfrac{(z_\nu - \zeta)}{(\zeta - x_\nu)}} = -\frac{Nf'(x_\nu)}{D(\zeta - x_\nu)S}. \tag{10.9}$$

4. Для доказательства сходимости последовательности $\{\mu_n\}$ представим μ_n в виде

$$\mu_n = \mu_0 \prod_{\nu=0}^{n-1} \frac{\mu_{\nu+1}}{\mu_\nu} \quad (\mu_\nu > 0,\ \nu = 0, 1, \ldots) \tag{10.10}$$

и воспользуемся двумя известными теоремами о сходимости бесконечных произведений.

Теорема А. Если ряд $\sum_{\nu=0}^{\infty} |c_\nu - 1|$ сходится и все c_ν отличны от нуля, то бесконечное произведение $\prod_{\nu=0}^{\infty} c_\nu$ сходится (т. е. $\prod_{\nu=0}^{n} c_\nu$ имеет конечный предел, отличный от нуля).

Теорема В. Если множители c_ν являются непрерывными функциями точки и ряд $\sum_{\nu=0}^{\infty} |c_\nu - 1|$ сходится равномерно в некоторой области, то бесконечное произведение $\prod_{\nu=0}^{\infty} c_\nu$ равномерно сходится к некоторой функции, непрерывной в этой области.

5. Учитывая (10.10) и (10.9), представим μ_n в виде отношения двух произведений

$$\mu_n = \frac{\rho_0 \prod_{\nu=0}^{n-1} Q_\nu}{\prod_{\nu=0}^{n-1} q_\nu}, \tag{10.11}$$

где

$$Q_\nu = -\frac{N}{(\zeta - x_\nu) S}; \quad q_\nu = \frac{D}{f'(x_\nu)} \quad (\nu = 0, 1, \ldots). \tag{10.12}$$

Докажем сходимость бесконечных произведений $\prod_{\nu=0}^{\infty} Q_\nu$ и $\prod_{\nu=0}^{\infty} q_\nu$ с помощью теоремы А.

Рассмотрим сначала

$$Q_\nu - 1 = \frac{1}{S}\left(-\frac{N}{\zeta - x_\nu} - S\right). \tag{10.13}$$

По формулам (10.6) и (10.7) имеем

$$-\frac{N}{\zeta - x_\nu} - S = \int_0^1 \int_0^1 \{f''[z_\nu + t(x_\nu - z_\nu) + ut(\zeta - x_\nu)] - \\ - f''[x_\nu + ut(\zeta - x_\nu)]\} t\, du\, dt.$$

Применяя (9.6), получаем

$$-\frac{N}{\zeta - x_\nu} - S = (z_\nu - x_\nu) \int_0^1 \int_0^1 \int_0^1 t(1-t) f''' dw\, du\, dt,$$

где $f''' = f'''[x_\nu + ut(\zeta - x_\nu) + w(1-t)(z_\nu - x_\nu)]$

Так как $t(1-t) \geqslant 0$ при $0 \leqslant t \leqslant 1$, то можно применить обобщенную теорему о среднем значении в интегральном исчислении:

$$-\frac{N}{\zeta - x_\nu} - S = (z_\nu - x_\nu) f'''(\xi) \int_0^1 \int_0^1 \int_0^1 t(1-t)\, dt\, du\, dw,$$

$$\xi \in (x_\nu, z_\nu).$$

Отсюда после интегрирования получаем

$$-\frac{N}{\zeta - x_\nu} - S = \frac{1}{6}(z_\nu - x_\nu) f'''(\xi), \quad \xi \in (x_\nu, z_\nu). \tag{10.14}$$

Применяя ту же теорему к интегралу (10.7), получим

$$S = f''(\xi_1) \int_0^1 \int_0^1 t\, dt\, du = \frac{1}{2} f''(\xi_1), \quad \xi_1 \in (x_\nu, z_\nu). \tag{10.15}$$

Из (10.13), (10.14) и (10.15) следует

$$Q_\nu - 1 = \frac{1}{3}(z_\nu - x_\nu) \frac{f'''(\xi)}{f''(\xi_1)}, \quad \xi, \xi_1 \in (x_\nu, z_\nu). \tag{10.16}$$

6. Теперь рассмотрим

$$q_\nu - 1 = \frac{D - f'(x_\nu)}{f'(x_\nu)}. \tag{10.17}$$

Из формулы (10.15), применяя (9.6), имеем

$$D - f'(x_\nu) = \int_0^1 \{f'[x_\nu + t(z_\nu - x_\nu)] - f'(x_\nu)\}\, dt =$$

$$= (z_\nu - x_\nu) \int_0^1 \int_0^1 f''[x_\nu + ut(z_\nu - x_\nu)]\, t\, du\, dt.$$

Снова применяя обобщенную теорему о среднем, получаем

$$D - f'(x_\nu) = (z_\nu - x_\nu) f''(\xi_2) \int_0^1 \int_0^1 t\, du\, dt =$$

$$= \frac{1}{2}(z_\nu - x_\nu) f''(\xi_2), \quad \xi_2 \in (x_\nu, z_\nu).$$

Таким образом,

$$q_\nu - 1 = \frac{1}{2}(z_\nu - x_\nu) \frac{f''(\xi_2)}{f'(x_\nu)}, \quad \xi_2 \in (x_\nu, z_\nu). \qquad (10.18)$$

7. Оценим теперь (10.16) и (10.18). Обозначим при фиксированном y_0 $(y_0 > \zeta)$:

$$m_k = \min_{[x_0,\, y_0]} |f^{(k)}(x)|, \quad M_k = \max_{[x_0,\, y_0]} |f^{(k)}(x)| \quad (k = 1, 2, 3),$$

$$K = \max\left(\frac{1}{3}\frac{M_3}{m_2}, \ \frac{1}{2}\frac{M_2}{m_1}\right).$$

Из (10.16) и (10.18) непосредственно следуют оценки

$$|Q_\nu - 1| \leqslant |z_\nu - x_\nu| K; \quad |q_\nu - 1| \leqslant |z_\nu - x_\nu| K. \qquad (10.19)$$

8. Нам остается теперь доказать сходимость ряда

$$\sum_{\nu=0}^{\infty} |z_\nu - x_\nu| K = K \sum_{\nu=0}^{\infty} |z_\nu - x_\nu|,$$

мажорирующего оба ряда

$$\sum_{\nu=0}^{\infty} |Q_\nu - 1| \quad \text{и} \quad \sum_{\nu=0}^{\infty} |q_\nu - 1|.$$

Так как

$$|z_\nu - x_\nu| = |(\zeta - x_\nu) - (\zeta - z_\nu)| \leqslant |\zeta - x_\nu| + |\zeta - z_\nu|,$$

то достаточно доказать сходимость рядов

$$\sum_{\nu=0}^{\infty} |\zeta - x_\nu| \quad \text{и} \quad \sum_{\nu=0}^{\infty} |\zeta - z_\nu|.$$

Сходимость этих последних рядов легко устанавливается по признаку Даламбера, так как из (9.13) следует

$$\frac{\zeta - x_{\nu+1}}{\zeta - x_\nu} \sim -\frac{1}{2}\frac{f''(\zeta)}{f'(\zeta)}(\zeta - x_\nu) \to 0,$$

а из (3.11) следует

$$\frac{\zeta - z_{\nu+1}}{\zeta - z_\nu} \sim -\frac{1}{2}\frac{f''(\zeta)}{f'(\zeta)}(\zeta - x_\nu) \to 0.$$

9. Докажем, наконец, что ряд $\sum_{\nu=0}^{\infty} |z_\nu - x_\nu|$ *сходится равномерно* относительно z_0 в интервале $\zeta < z_0 \leqslant y_0$ при фиксированном y_0. Заметим, что точки x_ν не зависят от z_0. Зафиксируем x_0 и исследуем поведение ряда $\sum_{\nu=0}^{\infty} |z_\nu - x_\nu|$ при $z_0 \downarrow \zeta$.

Все z_ν убывают с убыванием z_0. Возьмем $z_0' = y_0$ в качестве начального значения для последовательности (10.1); соответствующие значения z_ν обозначим через z_ν'. Тогда при любом $z_0 \leqslant y_0$ имеем

$$z_\nu \leqslant z_\nu' \quad (\nu = 0, 1, \ldots)$$

и, следовательно,

$$|z_\nu - x_\nu| \leqslant |z_\nu' - x_\nu|.$$

Из формул (10.19) мы получаем теперь оценки

$$|Q_\nu - 1| \leqslant K|z_\nu' - x_\nu|; \quad |q_\nu - 1| \leqslant K|z_\nu' - x_\nu|,$$

справедливые для всего интервала $\zeta < z_0 \leqslant y_0$. Поэтому ряды $\sum_{\nu=0}^{\infty} |Q_\nu - 1|$ и $\sum_{\nu=0}^{\infty} |q_\nu - 1|$ сходятся *равномерно* в этом интервале. По теореме В бесконечные произведе-

ния $\prod_{\nu=0}^{\infty} Q_\nu$ и $\prod_{\nu=0}^{\infty} q_\nu$ также сходятся равномерно в этом интервале и, значит, последовательность $\{\mu_n\}$ равномерно сходится к $\Lambda(z_0)$. Заметим теперь, что z_ν — непрерывные функции от z_0. По формуле (10.2) μ_ν — также непрерывные функции от z_0, следовательно, $\Lambda(z_0)$ — непрерывная функция от z_0 в $\zeta < z_0 \leqslant y_0$.

10. Если $z_0 = \zeta$, то все $z_\nu = \zeta$ $(\nu = 0, 1, \ldots)$; поэтому при $z_0 \downarrow \zeta$ все $\mu_\nu(z_0) \downarrow 0$. Так как μ_ν являются монотонными функциями от z_0, то $\Lambda(z_0)$ также монотонна на интервале $\zeta < z_0 \leqslant y_0$. Но $\Lambda(z_0) \geqslant 0$, поэтому $\Lambda(z_0) \to \Lambda_0$ при $z_0 \downarrow \zeta$. С другой стороны, для каждого положительного числа ε можно найти функцию $\mu_k(z_0)$, такую, что

$$|\Lambda(z_0) - \mu_k(z_0)| \leqslant \varepsilon \quad (\zeta < z_0 \leqslant y_0).$$

Поэтому

$$\Lambda(z_0) \leqslant \varepsilon + \mu_k(z_0);$$

переходя здесь к пределу при $z_0 \downarrow \zeta$, получаем $\Lambda_0 \leqslant \varepsilon$. Итак, $\Lambda_0 = 0$, т. е. $\Lambda(z_0) \downarrow 0$ при $z_0 \downarrow \zeta$. Мы видим, что отклонение приближения z_ν от корня может быть сделано произвольно малой частью отклонения приближения от корня x_ν, если начальная точка z_0 выбрана достаточно близко к корню.

ГЛАВА 11

ТРИ УЗЛА ИНТЕРПОЛЯЦИИ

Интерполяция дробно-линейными функциями

1. Пусть $f(x)$ определена на J_x. Пусть заданы *три различных узла интерполяции*

$$x_\nu \in J_x, \quad f(x_\nu) = f_\nu \quad (\nu = 0,\ 1,\ 2). \tag{11.1}$$

Построим функцию, интерполирующую $f(x)$ по этим трем узлам. Если бы в качестве такой функции мы выбрали многочлен, то он должен был бы быть по меньшей мере квадратичным; поэтому он имел бы по меньшей мере два корня, и мы не смогли бы узнать, какой из них выбрать. Этих трудностей можно избежать, если вместо многочлена рассмотреть функцию, имеющую лишь *один* простой корень, а именно дробно-линейную функцию

$$w = \frac{\alpha x + \beta}{\gamma x + \delta}. \tag{11.2}$$

Заметим, что в (11.2) мы имеем три существенные константы. Из проективной геометрии известно, что если w связано с x формулой (11.2), то точки x, x_0, x_1, x_2 имеют то же двойное отношение, что и w, f_0, f_1, f_2, т. е.

$$\frac{x - x_1}{x - x_0} : \frac{x_2 - x_1}{x_2 - x_0} = \frac{w - f_1}{w - f_0} : \frac{f_2 - f_1}{f_2 - f_0}. \tag{11.3}$$

Предположим, что числа f_0, f_1, f_2 различны, и введем Δ по формуле

$$\Delta = \frac{f_2 - f_1}{f_2 - f_0} : \frac{x_2 - x_1}{x_2 - x_0} \quad (f_0 \neq f_1 \neq f_2 \neq f_0, \quad x_0 \neq x_1 \neq x_2 \neq x_0). \tag{11.4}$$

Тогда

$$\frac{w - f_1}{w - f_0} = \Delta\, \frac{x - x_1}{x - x_0} \quad (f_0 \neq f_1 \neq f_2 \neq f_0, \quad x_0 \neq x_1 \neq x_2 \neq x_0). \tag{11.5}$$

Таким способом мы строим интерполирующую функцию w, когда все три узла интерполяции различны.

Два совпадающих узла интерполяции

2. В дальнейшем мы будем рассматривать случай, когда две из трех точек x_ν совпадают, т. е. $x_2 = x_0$, $x_1 \neq x_0$. Мы предполагаем, что известны f_0, f_1 и $f_0' = f'(x_0)$ и что $f_0 \neq f_1$, $f_0' \neq 0$. Интерполирующую функцию в этом случае получим, переходя к пределу в (11.4) и (11.5) при $x_2 \to x_0$.

Перепишем (11.4) в виде

$$\Delta = \frac{f_2 - f_1}{x_2 - x_1} : \frac{f_2 - f_0}{x_2 - x_0}. \tag{11.6}$$

При $x_2 \to x_0$ пределом Δ будет

$$\Delta^* = \frac{1}{f_0'} \frac{f_1 - f_0}{x_1 - x_0} \quad (f_0 \neq f_1, \quad x_0 \neq x_1). \tag{11.7}$$

Формула (11.5) примет вид

$$\frac{w - f_1}{w - f_0} = \Delta^* \frac{x - x_1}{x - x_0}. \tag{11.8}$$

Непосредственно из формулы (11.8) заключаем, что w удовлетворяет условиям $w(x_0) = f_0$ и $w(x_1) = f_1$. Дифференцируя обе части (11.8) по x, можно проверить, что также и $w'(x_0) = f_0'$.

3. Как и в предыдущих исследованиях, для оценки погрешности мы воспользуемся обратной функцией. Пусть $x = \Phi(w)$ есть функция, обратная к $w = f(x)$. Тогда

$$\Phi(f_0) = x_0, \quad \Phi(f_1) = x_1, \quad \Phi'(f_0) = \frac{1}{f_0'}.$$

С другой стороны, уравнение (11.8) определяет интерполирующую функцию $x = \varphi(w)$. Явное выражение этой функции мы получим, разрешая уравнение

$$\frac{w - f_1}{w - f_0} = \Delta^* \frac{\varphi - x_1}{\varphi - x_0} \tag{11.9}$$

относительно φ:

$$\varphi(w) = \frac{(x_0 - x_1 \Delta^*) w + x_1 f_0 \Delta^* - x_0 f_1}{(1 - \Delta^*) w + f_0 \Delta^* - f_1}. \tag{11.10}$$

Для оценки погрешности нам понадобится производная третьего порядка φ'''. Так как при любых постоянных α, β, γ, δ

$$\left(\frac{\alpha w+\beta}{\gamma w+\delta}\right)''' = 6\gamma^2 \frac{\alpha\delta-\beta\gamma}{(\gamma w+\delta)^4}, \tag{11.11}$$

то

$$\varphi'''(w) = 6(1-\Delta^*)^2 \frac{\Delta^*(f_1-f_0)(x_1-x_0)}{N^4}, \tag{11.12}$$

где

$$N = (1-\Delta^*)\,w + f_0\,\Delta^* - f_1. \tag{11.13}$$

Оценка погрешности

4. Пусть $\varphi(w) \supset \Phi(w)$, так что в качестве нового приближения к корню $\zeta = \Phi(0)$ мы берем $\varphi(0) = x_2$.

Значение x_2 получается из (11.10) при $w = 0$

$$x_2 = \frac{x_1 f_0 \Delta^* - x_0 f_1}{f_0 \Delta^* - f_1}. \tag{11.14}$$

Для оценки погрешности этого приближения воспользуемся формулой (1.4а), которая в рассматриваемом случае имеет вид

$$\Phi(f) - \varphi(f) = \frac{1}{6}(f-f_0)^2(f-f_1)\,[\Phi'''(\eta) - \varphi'''(\eta)],$$
$$\eta \in (f,\ f_0,\ f_1). \tag{11.15}$$

Полагая здесь $f = 0$, получаем

$$\zeta - x_2 = -\frac{f_0^2 f_1}{6}\,[\Phi'''(\eta) - \varphi'''(\eta)], \quad \eta \in (0,\ f_0,\ f_1). \tag{11.16}$$

Исследуем порядок величины правой части (11.16) при $x_0, x_1 \to \zeta$. При этом $f_0, f_1 \to f(\zeta) = 0$ и, значит, $\eta \to 0$. Далее, из формулы (11.7) следует, что

$$\Delta^* = \frac{1}{f'(x_0)}\,\frac{f(x_1)-f(x_0)}{x_1-x_0} \to \frac{f'(\zeta)}{f'(\zeta)} = 1, \tag{11.17}$$

если только $f'(\zeta) \neq 0$.

Рассмотрим отношение

$$\frac{\Delta^*-1}{x_1-x_0} = \frac{f(x_1) - [f(x_0) + f'(x_0)(x_1-x_0)]}{f'(x_0)(x_1-x_0)^2}. \tag{11.18}$$

Здесь выражение в квадратных скобках дает первые два члена разложения Тейлора функции $f(x_1)$ в окрестности точки x_0. Поэтому числитель правой части (11.18) равен остаточному члену $\frac{1}{2}(x_1-x_0)^2 f''(\xi)$, и мы имеем

$$\frac{\Delta^*-1}{x_1-x_0}=\frac{1}{2}\frac{f''(\xi)}{f'(x_0)}, \quad \xi\in(x_0,\ x_1), \tag{11.19}$$

откуда при $x_0, x_1 \to \zeta$ следует

$$\frac{\Delta^*-1}{x_1-x_0}\to\frac{1}{2}\frac{f''(\zeta)}{f'(\zeta)}. \tag{11.20}$$

Рассмотрим, далее, отношение

$$\frac{N}{x_1-x_0}=-\frac{\Delta^*-1}{x_1-x_0}w+f_0\frac{\Delta^*-1}{x_1-x_0}-\frac{f_1-f_0}{x_1-x_0}. \tag{11.21}$$

При x_0, $x_1\to\zeta$ и $w\to 0$, учитывая (11.20), получаем

$$\frac{N}{x_1-x_0}\to -f'(\zeta). \tag{11.22}$$

5. Теперь из формул (11.12), (11.17), (11.20) и (11.22) при x_0, $x_1\to\zeta$ находим

$$\varphi'''(\eta)=6\Delta^*\left(\frac{\Delta^*-1}{x_1-x_0}\right)^2\left(\frac{f_1-f_0}{x_1-x_0}\right)\left(\frac{x_1-x_0}{N}\right)^4\to$$

$$\to 6\cdot 1\left[\frac{f''(\zeta)}{2f'(\zeta)}\right]^2 f'(\zeta)\left[\frac{1}{f'(\zeta)}\right]^4,$$

$$\varphi'''(\eta)\to\frac{3}{2}\frac{f''(\zeta)^2}{f'(\zeta)^5}. \tag{11.23}$$

Далее, из формулы (2.5) и таблицы п. 8 гл. 2 имеем

$$\Phi'''(\eta)=\frac{3(f'')^2-f'f'''}{(f')^5}. \tag{11.24}$$

Если $\eta\to 0$, то $\Phi(\eta)\to\zeta$, поэтому

$$\Phi'''(\eta)\to\frac{3f''(\zeta)^2-f'(\zeta)f'''(\zeta)}{f'(\zeta)^5}. \tag{11.25}$$

Таким образом,

$$\Phi'''(\eta)-\varphi'''(\eta)\to\frac{3f''(\zeta)^2-2f'(\zeta)f'''(\zeta)}{2f'(\zeta)^5}\quad (x_0, x_1\to\zeta). \tag{11.26}$$

Остается оценить f_0 и f_1

$$f_0 = f(x_0) - f(\zeta) = f'(\xi_0)(x_0 - \zeta), \quad \xi_0 \in (x_0, \zeta), \quad (11.27)$$
$$f_1 = f(x_1) - f(\zeta) = f'(\xi_1)(x_1 - \zeta), \quad \xi_1 \in (x_1, \zeta). \quad (11.28)$$

Подставляя эти результаты в формулу (11.16) и учитывая, что $f'(\xi_0)$, $f'(\xi_1) \to f'(\zeta)$ при $x_0, x_1 \to \zeta$, получаем

$$\frac{\zeta - x_2}{(\zeta - x_0)^2(\zeta - x_1)} \to \frac{3f''(\zeta)^2 - 2f'(\zeta)f'''(\zeta)}{12f'(\zeta)^2}. \quad (11.29)$$

Формула (11.29) показывает, что в данном случае мы имеем *аппроксимацию третьего порядка*.

Применение в итерационном процессе

6. Рассматриваемый метод интересно сравнить с комбинацией метода Ньютона и regula falsi.

Если мы уже применили оба метода, метод Ньютона в точке x_0 и regula falsi для точек x_0 и x_1, то без затраты дополнительных горнеров мы сразу получаем по формуле (11.14) результат, существенно лучший, чем результаты, полученные с помощью обоих этих методов. Конечно, такое улучшение результата (без затраты дополнительных горнеров) возможно только *один раз*.

С другой стороны, если мы хотим применять этот метод аппроксимации как итерационный, то мы можем последовательно использовать следующие тройки чисел:

$$(x_0, x_1, x_1), \quad (x_1, x_2, x_2), \quad (x_2, x_3, x_3), \ldots,$$

где в первой тройке x_1 используется *дважды*, а x_0 только *один раз*.

7. Для определения индекса эффективности рассматриваемого метода положим

$$\ln \frac{1}{|x_\mu - \zeta|} = y_\mu. \quad (11.30)$$

Из (11.29), поменяв местами x_0 и x_1, получим

$$y_{\mu+2} = y_\mu + 2y_{\mu+1} + k_\mu, \quad (11.31)$$

где k_μ ограничены, если правая часть (11.29) отлична от

нуля. В следующей главе мы покажем, что если $x_\mu \to \zeta$, $y_\mu \to \infty$, то отношение $y_{\mu+1} : y_\mu$ стремится к пределу

$$\frac{y_{\mu+1}}{y_\mu} \to 1 + \sqrt{2} = 2{,}414\ldots \qquad (11.32)$$

Так как на каждом шаге расчета мы затрачиваем два горнера, то индекс эффективности рассматриваемого метода равен $\sqrt{2{,}414\ldots} = 1{,}55\ldots$. Сравнивая его с индексами эффективности метода секущих (1,618...) и метода Ньютона (1,414...), заключаем, что этот новый метод итерации лучше метода Ньютона, но хуже метода секущих.

Пример. Применим данный метод к уравнению $x^3 - 2x - 5 = 0$, исследованному в п. 17 гл. 3. Начиная с тройки (x_0, x_1, x_1), $x_0 = 3$, $x_1 = 2$, получим следующую последовательность чисел [1]):

ν	x_ν	$\lvert x_\nu - \zeta \rvert$	$\frac{y_{\nu+1}}{y_\nu}$
2	2,0960451977401129943	$1{,}4937 \cdot 10^{-3}$	2,759
3	2,0945514320381108026	$4{,}95 \cdot 10^{-8}$	2,585
4	2,0945514815423265923	$9 \cdot 10^{-19}$	2,47
5	2,0945514815423265914	—	—

[1]) Вычисления проведены А. Рейтером в Математическом исследовательском центре США, Мэдисон, штат Висконсин.

ГЛАВА 12

ЛИНЕЙНЫЕ РАЗНОСТНЫЕ УРАВНЕНИЯ

Неоднородные и однородные разностные уравнения

1. В этой главе мы будем изучать так называемые *линейные разностные уравнения n-го порядка с постоянными коэффициентами*, т. е. некоторую последовательность рекуррентных формул

$$a_0 z_{\mu+n} + a_1 z_{\mu+n-1} + \ldots + a_n z_\mu = k_{\mu+n} \quad (\mu = 0, 1, \ldots), \tag{12.1}$$

где $a_0 = 1$, $a_1, \ldots, a_n$ $(a_n \neq 0)$ — фиксированные постоянные, а $k_{\mu+n}$ — заданная последовательность. Задача обычно состоит в том, чтобы определить все z_μ $(\mu \geqslant 0)$, если заданы начальные значения $z_0, z_1, \ldots, z_{n-1}$. Это, конечно, всегда возможно.

Если все $k_\mu = 0$, то мы имеем *однородное разностное уравнение*:

$$a_0 z_{\mu+n} + a_1 z_{\mu+n-1} + \ldots + a_n z_\mu = 0$$
$$(a_0 = 1;\ \mu = 0,\ 1, \ldots). \tag{12.2}$$

2. Решения уравнений (12.1) и (12.2) зависят от так называемого *характеристического многочлена* этих уравнений

$$\varphi(x) = x^n + a_1 x^{n-1} + \ldots + a_n. \tag{12.3}$$

Одновременно с $\varphi(x)$ мы будем рассматривать многочлен

$$\psi(x) = x^n \varphi\left(\frac{1}{x}\right) = 1 + a_1 x + \ldots + a_n x^n. \tag{12.4}$$

Если постоянные k_ν заданы, то можно считать, что задан степенной ряд

$$K(x) = \sum_{\nu=n}^{\infty} k_\nu x^\nu, \tag{12.5}$$

и можно рассматривать задачу определения z_ν из (12.1) как задачу определения соответствующего *производящего ряда*

$$Z(x) = \sum_{\nu=0}^{\infty} z_\nu x^\nu. \tag{12.6}$$

3. Если мы умножим $Z(x)$ на $\psi(x)$, то в полученном произведении коэффициенты при x^ν для $\nu \geqslant n$ будут вследствие (12.1) совпадать с соответствующими k_ν. Следовательно, (12.1) эквивалентно уравнению

$$\psi(x) Z(x) = K(x) + P_{n-1}(x), \tag{12.7}$$

где $P_{n-1}(x)$ — полином степени $n-1$, который может быть выбран произвольно.

Разрешая (12.7) относительно $Z(x)$, получаем

$$Z(x) = \frac{K(x)}{\psi(x)} + \frac{P_{n-1}(x)}{\psi(x)}. \tag{12.8}$$

Таким образом, z_ν могут быть получены разложением правой части (12.8) в ряд по степеням x.

Общее решение однородного уравнения

4. В случае однородного уравнения (12.2) функция $K(x)$ равна 0 и разложение $Z(x)$ в ряд можно получить, например, используя разложение $\frac{P_{n-1}(x)}{\psi(x)}$ на простейшие дроби.

Обозначим n корней характеристического многочлена $\varphi(x)$ через $u_1, \ldots, u_n$ и перенумеруем их так, чтобы выполнялось условие

$$|u_1| > |u_2| \geqslant \ldots \geqslant |u_n|.$$

Предположим сначала, что все u_ν различны. В этом случае правая часть выражения (12.8) при $K(x) = 0$ равна

$$\sum_{k=1}^{n} \frac{\gamma_k}{1 - u_k x},$$

и мы получаем

$$z_\mu = \sum_{k=1}^{n} \gamma_k u_k^\mu \quad (\mu = 0, 1, \ldots). \tag{12.9}$$

Здесь n постоянных γ_k — «постоянные интегрирования» уравнения (12.2) — могут быть выбраны произвольно.

5. Если $\varphi(x)$ имеет кратные корни и, например, u является корнем кратности $m > 1$, то соответствующая часть разложения $\frac{P_{n-1}(x)}{\psi(x)}$ на простейшие дроби такова:

$$\frac{\gamma_1}{1-ux} + \frac{\gamma_2}{(1-ux)^2} + \ldots + \frac{\gamma_m}{(1-ux)^m}.$$

Разлагая здесь каждое слагаемое в степенной (биномиальный) ряд и собирая члены с одинаковыми степенями x, получаем

$$\sum_{\nu=0}^{\infty} u^\nu (b_0 + b_1 \nu + \ldots + b_{m-1} \nu^{m-1}) x^\nu,$$

где коэффициенты $b_0, b_1, \ldots, b_{m-1}$ можно считать произвольными постоянными, если $\gamma_1, \gamma_2, \ldots, \gamma_m$ подобраны соответствующим образом [1]). Другими словами, в этом случае

---

[1]) Действительно, при целом положительном k разложение $(1-ux)^{-k}$ в биномиальный ряд имеет вид

$$(1-ux)^{-k} = \sum_{\nu=0}^{\infty} \binom{-k}{\nu} (-ux)^\nu =$$

$$= 1 + \sum_{\nu=1}^{\infty} \frac{k(k+1)\ldots(k+\nu-1)}{\nu!} u^\nu x^\nu =$$

$$= \sum_{\nu=0}^{\infty} \frac{(\nu+k-1)\ldots(\nu+1)}{(k-1)!} u^\nu x^\nu =$$

$$= \sum_{\nu=0}^{\infty} \left(C_0^{(k)} + C_1^{(k)} \nu + \ldots + C_{k-1}^{(k)} \nu^{k-1}\right) u^\nu x^\nu,$$

где коэффициенты $C_0^{(k)}, C_1^{(k)}, \ldots, C_{k-1}^{(k)}$ не зависят от ν. Умножая это разложение на γ_k и суммируя по k от 1 до m, получаем

$$b_0 = \gamma_1 C_0^{(1)} + \gamma_2 C_0^{(2)} + \ldots + \gamma_m C_0^{(m)}; \ldots;$$

$$b_{m-2} = \gamma_{m-1} C_{m-2}^{(m-1)} + \gamma_m C_{m-2}^{(m)}; \quad b_{m-1} = \gamma_m C_{m-1}^{(m)},$$

откуда непосредственно следует утверждение автора. — *Прим. перев.*

«постоянными интегрирования» можно считать $b_0, b_1, \ldots, b_{m-1}$. Поступая аналогичным образом со всеми корнями $\varphi(x)$, мы получим следующее выражение для общего решения (12.2):

$$z_\mu = u^{(1)\mu} Q_1(\mu) + u^{(2)\mu} Q_2(\mu) + \ldots + u^{(k)\mu} Q_k(\mu), \qquad (12.10)$$

где $u^{(1)}, u^{(2)}, \ldots, u^{(k)}$ — *различные* корни $\varphi(x)$, имеющие кратности соответственно $m_1, m_2, \ldots, m_k$; $Q_1(\mu), Q_2(\mu), \ldots, Q_k(\mu)$ — произвольные многочлены относительно μ степени $m_1 - 1, m_2 - 1, \ldots, m_k - 1$ соответственно.

Лемма о делении степенных рядов

6. В дальнейшем мы воспользуемся формулой (12.8) для исследования асимптотического поведения z_ν в нескольких важных случаях.

При исследовании будет очень полезна

Лемма 12.1. Допустим, что при некоторых положительных постоянных s и γ коэффициенты степенного ряда

$$K(x) = \sum_{\nu=0}^{\infty} k_\nu x^\nu$$

удовлетворяют условию $|k_\nu| \leqslant \gamma s^\nu$ $(\nu = 0, 1, \ldots)$. Тогда при $|\xi| < s$ коэффициенты разложения

$$\frac{K(x)}{1 - \xi x} = \sum_{\nu=0}^{\infty} l_\nu x^\nu$$

удовлетворяют неравенству $|l_\nu| \leqslant \gamma_1 s^\nu$, где $\gamma_1 = \frac{\gamma s}{|s - |\xi||}$. То же самое неравенство справедливо и при $|\xi| > s$, если только $K\left(\frac{1}{\xi}\right) = 0$.

Доказательство. Предположим без потери общности, что $s = 1$, так как в противном случае можно заменить x на $\frac{x}{s}$. Если $|\xi| < 1$, то из равенства

$$l_\nu = k_\nu + k_{\nu-1}\xi + \ldots + k_0 \xi^\nu$$

следует

$$|l_\nu| \leqslant \gamma(1 + |\xi| + \ldots + |\xi|^\nu) \leqslant \frac{\gamma}{1 - |\xi|}.$$

Если $|\xi| > 1$ и $K\left(\frac{1}{\xi}\right) = 0$, то

$$l_\nu = \xi^\nu \left(k_0 + \frac{k_1}{\xi} + \ldots + \frac{k_\nu}{\xi^\nu}\right) = -\xi^\nu \left(\frac{k_{\nu+1}}{\xi^{\nu+1}} + \frac{k_{\nu+2}}{\xi^{\nu+2}} + \ldots\right),$$

$$|l_\nu| \leqslant \gamma\left(\frac{1}{|\xi|} + \frac{1}{|\xi|^2} + \ldots\right) = \frac{\gamma}{|\xi| - 1}.$$

7. Следствие 1. Предположим, что $K(x)$ удовлетворяет условиям леммы 12.1 и что даны n чисел $\xi_1, \xi_2, \ldots \xi_n$, не равных нулю и отличных по модулю от s. Предположим далее, что для каждого ξ_ν с $|\xi_\nu| > s$ функция $K(x)$ имеет корень $\frac{1}{\xi_\nu}$, кратность которого по меньшей мере равна количеству ξ_k, которые совпадают с ξ_ν. Тогда в разложении

$$\frac{K(x)}{\prod\limits_{\nu=1}^{n}(1 - \xi_\nu x)} = \sum_{\nu=0}^{\infty} t_\nu x^\nu$$

коэффициенты t_ν удовлетворяют условию $t_\nu = O(s^\nu)$.

Следствие 2. Если $|\xi| < s$, то разложение функции $\frac{1}{(1 - \xi x)(1 - sx)^m}$ при целом положительном m мажорируется разложением функции $\frac{\gamma_1}{(1 - sx)^m}$ с $\gamma_1 = \frac{s}{s - |\xi|}$:

$$\frac{1}{(1 - \xi x)(1 - sx)^m} \ll \frac{s}{s - |\xi|} \cdot \frac{1}{(1 - sx)^m} \quad (m = 1, 2, \ldots)\ ^{1)}.$$

---

[1]) Символ $\ll$ есть знак мажорирования, введенный Анри Пуанкаре. Он означает, что для любого значения ν модуль коэффициента при x^ν в левой части не превосходит соответствующего коэффициента в правой части.

Действительно, при $m = 1$ это утверждение может быть записано в виде

$$\frac{(1-sx)^{-1}}{1-\xi x} \ll \gamma_1 (1-sx)^{-1}$$

и следует из первого утверждения леммы при $\gamma = 1$.

С другой стороны, отношение мажорирования, очевидно, сохраняется, если обе части умножить на ряд с положительными коэффициентами. Умножая обе части нашего соотношения на разложение в ряд дроби $\frac{1}{(1-sx)^{m-1}}$ (которое имеет положительные коэффициенты), получаем утверждение следствия 2.

Асимптотическое поведение решений уравнения (12.1)

8. Теорема 12.1. Предположим, что корни характеристического многочлена $\varphi(x)$ удовлетворяют условию

$$|u_1| > 1 > |u_2| \tag{12.11}$$

и что

$$k_\nu = O(s^\nu) \quad (\nu \to \infty), \tag{12.12}$$

где s — положительная постоянная, $s \leqslant 1$.

Тогда отношение $\frac{z_\nu}{u_1^\nu}$ имеет конечный предел:

$$\frac{z^\nu}{u_1^\nu} \to \alpha \quad (\nu \to \infty). \tag{12.13}$$

Кроме того, если $s > |u_2|$, то

$$z_\nu = \alpha u_1^\nu + O(s^\nu) \quad (\nu \to \infty). \tag{12.14}$$

Если $s = |u_2|$ и m — наибольшая кратность корней многочлена $\varphi(x)$, имеющих модуль, равный $|u_2|$, то

$$z_\nu = \alpha u_1^\nu + O\left(\nu^m u_2^\nu\right) \quad (\nu \to \infty). \tag{12.15}$$

Наконец, если $s < |u_2|$, то (при том же значении m)

$$z_\nu = \alpha u_1^\nu + O\left(\nu^{m-1} u_2^\nu\right) \quad (\nu \to \infty). \tag{12.16}$$

9. Доказательство. Пусть $l - 1$ корней имеют модуль, равный $|u_2|$:

$$|u_2| = \ldots = |u_l| > |u_{l+1}|.$$

Без ограничения общности можно считать, что $s > |u_{l+1}|$.

Воспользовавшись следствием 1 из леммы 12.1, заключаем, что правая часть (12.8) может быть представлена в виде

$$Z(x) = \frac{K_1(x)}{\prod\limits_{\lambda=1}^{l}(1-u_\lambda x)}, \tag{12.17}$$

где

$$K_1(x) = \sum_{\nu=0}^{\infty} k'_\nu x^\nu, \quad k'_\nu = O(s^\nu) \quad (\nu \to \infty).$$

Так как ряд для $K_1(x)$ имеет радиус сходимости $\geqslant \frac{1}{s} > \frac{1}{|u_1|}$, то $Z(x)$ имеет простой полюс в точке $\frac{1}{u_1}$. Поэтому

$$Z(x) = \frac{\alpha}{1-u_1 x} + P(x), \tag{12.18}$$

где $P(x)$ — ряд, имеющий радиус сходимости $> \frac{1}{|u_1|}$.

Из формул (12.17) и (12.18) получаем далее

$$P(x) = \frac{K_2(x)}{(1-u_1 x)\prod\limits_{\lambda=2}^{l}(1-u_\lambda x)}, \tag{12.19}$$

$$K_2(x) = \sum_{\nu=0}^{\infty} k''_\nu x^\nu = K_1(x) - \alpha \prod_{\lambda=2}^{l}(1-u_\lambda x), \tag{12.20}$$

где $k''_\nu = O(s^\nu)$. Очевидно, что $K_2\left(\frac{1}{u_1}\right) = 0$. По лемме 12.1 для $\xi = u_1$ коэффициент при x^ν ряда

$$\frac{K_2(x)}{1-u_1 x} = K_3(x) = \sum_{\nu=0}^{\infty} k'''_\nu x^\nu$$

является величиной $O(s^\nu)$.

Из формул (12.18) и (12.19) следует

$$Z(x) - \frac{\alpha}{1-u_1 x} = \frac{K_3(x)}{\prod\limits_{\lambda=2}^{l}(1-u_\lambda x)}. \tag{12.21}$$

10. Если теперь $s > |u_2|$, то к правой части (12.21) применимо следствие 1 из леммы 12.1, поэтому ν-й коэффициент разложения правой части (12.21) в ряд есть $O(s^\nu)$. Отсюда непосредственно следуют формулы (12.13) и (12.14).

Рассмотрим теперь случай $s \leqslant |u_2|$. Пусть среди корней $u_2, \ldots, u_l$ имеется t различных. Обозначим их через $u^{(1)}, \ldots, u^{(t)}$, а их кратности соответственно через $m_1, \ldots, m_t$. Очевидно, что $m = \max\limits_{\tau=1,\ldots,t} m_\tau$.

Запишем разложение на простейшие дроби

$$\frac{1}{\prod\limits_{\lambda=2}^{l}(1-u_\lambda x)} = \sum_{\tau=1}^{t} \frac{p_\tau(x)}{(1-u^{(\tau)}x)^{m_\tau}},$$

где $p_\tau(x)$ — многочлен степени $m_\tau - 1$.

Подставим это разложение в формулу (12.21):

$$Z(x) - \frac{\alpha}{1-u_1 x} = \sum_{\tau=1}^{t} \frac{K^{(\tau)}(x)}{(1-u^{(\tau)}x)^{m_\tau}}. \qquad (12.22)$$

Здесь

$$K^{(\tau)}(x) = \sum_{\nu=0}^{\infty} k_{\tau,\nu} x^\nu = K_3(x)\, p_\tau(x) \qquad (\tau = 1, \ldots, t);$$

поэтому

$$k_{\tau,\nu} = O(s^\nu) \qquad (\nu \to \infty; \quad \tau = 1, \ldots, t),$$

и все $K^{(\tau)}(x)$ допускают мажоранту вида $\gamma^{(\tau)}(1-sx)^{-1}$. Следовательно, при соответствующем выборе γ разложение правой части (12.22) допускает мажоранту вида $\gamma(1-sx)^{-1}(1-|u_2|x)^{-m}$:

$$Z(x) - \frac{\alpha}{1-u_1 x} \ll \frac{\gamma}{(1-sx)(1-|u_2|x)^m}. \qquad (12.23)$$

Если при этом $s = |u_2|$, то коэффициент при x^ν в разложении правой части (12.23) равен $\gamma\binom{m+\nu}{\nu}|u_2|^\nu$, откуда следуют (12.13), (12.15).

Если же $s < |u_2|$, то по следствию 2 из леммы 12.1 (где ξ заменено на s и s на $|u_2|$) правая часть (12.23)

допускает мажоранту $\gamma_1(1-|u_2|x)^{-m}$. Поэтому коэффициент при x^ν в разложении $Z(x)-\frac{a}{1-u_1 x}$ есть $O(\nu^{m-1}|u_2|^\nu)$ при $\nu\to\infty$. Таким образом, справедливы (12.13) и (12.16). Теорема доказана.

Асимптотическое поведение ошибок при итерациях по методу секущих

11. Применим теперь нашу теорему для доказательства (3.19). Введем обозначения

$$\delta_\nu=|x_\nu-\zeta|,\quad y_\nu=\ln\frac{1}{\delta_\nu}. \tag{12.24}$$

Из формулы (3.13), заменяя x_0 на $x_{\nu-1}$, получим

$$y_{\nu+1}-y_\nu-y_{\nu-1}=\ln\left|\frac{2f'(\xi)^3}{f''(\xi)f'(\xi_1)f'(\xi_2)}\right|, \tag{12.25}$$

где ξ, ξ_1, ξ_2 лежат в наименьшем интервале, содержащем ζ, $x_{\nu-1}$, x_ν. Поэтому

$$\max(|\xi-\zeta|,\ |\xi_1-\zeta|,\ |\xi_2-\zeta|)\leqslant\max(\delta_{\nu-1},\ \delta_\nu). \tag{12.26}$$

Выражение в правой части формулы (12.25) при $\nu\to\infty$ имеет предел

$$\beta=\ln\left|\frac{2f'(\zeta)}{f''(\zeta)}\right|. \tag{12.27}$$

Мы предполагаем, конечно, что $f'(\zeta)f''(\zeta)\neq 0$ и что все x_ν лежат внутри интервала, в котором $f'(x)$ и $f''(x)$ не меняют знака. Правую часть формулы (12.25) положим равной $\beta+k_{\nu+1}$, где

$$k_{\nu+1}=3\ln\frac{f'(\xi)}{f'(\zeta)}-\ln\frac{f'(\xi_1)}{f'(\zeta)}-\ln\frac{f'(\xi_2)}{f'(\zeta)}-\ln\frac{f''(\xi)}{f''(\zeta)}.$$

12. Пусть в рассматриваемом интервале

$$\max|f''(x)|=M_2,\quad \max|f'''(x)|=M_3;$$
$$\min|f'(x)|=m_1,\quad \min|f''(x)|=m_2.$$

Из теоремы о среднем значении и из неравенства (12.26) заключаем, что

$$\max\left(\left|\ln\frac{f'(\xi)}{f'(\zeta)}\right|,\ \left|\ln\frac{f'(\xi_1)}{f'(\zeta)}\right|,\ \left|\ln\frac{f'(\xi_2)}{f'(\zeta)}\right|\right)\leqslant\frac{M_2}{m_1}\max(\delta_{\nu-1},\ \delta_\nu),$$

$$\left|\ln\frac{f''(\xi)}{f''(\zeta)}\right|\leqslant\frac{M_3}{m_2}\max(\delta_{\nu-1},\ \delta_\nu)\ ^{1)}.$$

Поэтому

$$k_{\nu+1}=O(\delta_{\nu-1}+\delta_\nu).$$

С другой стороны, из оценок, установленных в п. 9 и п. 10 гл. 3, следует, что

$$\delta_\nu=\left(O\,d^{t_1^{\nu+1}/\sqrt{5}}\right),$$

где $0<d<1$ и $t_1>1$. Очевидно, что $\sqrt[\nu]{\delta_\nu}\to 0$; поэтому ряд $\sum\limits_{\nu=0}^{\infty}k_\nu x^\nu$ определяет целую функцию. Полагая в формуле (12.25) $y_\nu=v_\nu-\beta$, получаем окончательно разностное уравнение

$$v_{\nu+1}-v_\nu-v_{\nu-1}=k_{\nu+1},\tag{12.28}$$

где $k_\nu=O(s^\nu)$ при любом сколь угодно малом $s>0$. Так как, кроме того, корни t_1, t_2 характеристического многочлена x^2-x-1 удовлетворяют неравенству $t_1>1>|t_2|$, то выполнены все условия теоремы 12.1. Следовательно,

$$y_\nu+\beta-\alpha t_1^\nu=v_\nu-\alpha t_1^\nu=O\left(t_2^\nu\right),\tag{12.29}$$

где α есть некоторая положительная постоянная, так как $y_\nu\to+\infty$ при $x_\nu\to\zeta$.

Из формул (12.24), (12.27) и (12.28) получаем

$$|x_\nu-\zeta|=\left|\frac{2f'(\zeta)}{f''(\zeta)}\right|e^{-\alpha t_1^\nu}\left[1+O\left(t_2^\nu\right)\right].\tag{12.30}$$

Возведем обе части этого равенства в степень t_1, затем перепишем (12.30) для $\nu+1$ и разделим полученные выражения для $|x_{\nu+1}-\zeta|$ и $|x_\nu-\zeta|^{t_1}$:

$$\frac{|x_{\nu+1}-\zeta|}{|x_\nu-\zeta|^{t_1}}=\left|\frac{2f'(\zeta)}{f''(\zeta)}\right|^{t_2}+O\left(t_2^\nu\right);\tag{12.31}$$

отсюда непосредственно следует (3.19).

---

[1]) Здесь теорема о среднем значении применяется к $\ln|f'(x)|$ и к $\ln|f''(x)|$, например,

$$\ln\frac{f'(\xi)}{f'(\zeta)}=\ln|f'(\xi)|-\ln|f'(\zeta)|=\frac{f''(c)}{f'(c)}(\xi-\zeta),\quad c\in(\xi,\ \zeta).$$

В гл. 11 мы получили разностное уравнение (11.31), в котором k_μ ограничены и характеристический многочлен которого x^2-2x-1 имеет два корня, $1+\sqrt{2}>1$ и $1-\sqrt{2}$ с модулем $\sqrt{2}-1<1$. Применяя и здесь теорему 12.1 при $k_\mu=O(1)$, мы получим

$$y_\mu=\alpha(1+\sqrt{2})^\mu+O(1),$$

где α — некоторая положительная постоянная, так как $y_\mu\to+\infty$ при $\mu\to\infty$. Отсюда непосредственно вытекает (11.32).

Теорема о корнях уравнений некоторого класса

13. В приложениях рассматриваемой теории часто бывает полезной следующая теорема Коши:

Если в уравнении

$$x^n-b_1x^{n-1}-\ldots-b_n=0 \qquad (12.32$$

все $b_\nu\geqslant 0$ $(\nu=1,2,\ldots,n)$, *но не все* b_ν *равны нулю, то уравнение* (12.32) *имеет единственный положительный корень* p *и модули всех остальных корней этого уравнения не превосходят* p.

Мы покажем, что справедливо более точное утверждение, имеющее большое значение для приложений.

Теорема 12.2. Если в уравнении (12.32) все $b_\nu\geqslant 0$, а индексы положительных b_ν имеют наибольшим общим делителем единицу, то уравнение (12.32) имеет единственный простой положительный корень p и модули всех остальных корней этого уравнения *меньше* p.

Доказательство. Пусть среди чисел $b_1,\ldots,b_n$ *положительны только* $b_{k_1}, b_{k_2},\ldots,b_{k_m}$, $k_1<k_2<\ldots<k_m\leqslant n$. По условию относительно индексов k_ν существует m целых чисел $s_1, s_2,\ldots,s_m$, таких, что

$$s_1k_1+s_2k_2+\ldots+s_mk_m=1. \qquad (12.33)$$

Уравнение (12.32) при $x\neq 0$ можно записать в виде

$$F(x)\equiv\frac{b_{k_1}}{x^{k_1}}+\frac{b_{k_2}}{x^{k_2}}+\ldots+\frac{b_{k_m}}{x^{k_m}}-1=0. \qquad (12.34)$$

Функция $F(x)$ при положительных x строго убывает от $+\infty$ до -1 и поэтому обращается в нуль только при одном положительном значении $x=p$, $p>0$. Ее производная в точке p есть

$$F'(p) = -k_1 \frac{b_{k_1}}{p^{k_1+1}} - k_2 \frac{b_{k_2}}{p^{k_2+1}} - \ldots - k_m \frac{b_{k_m}}{p^{k_m+1}} < 0,$$

поэтому p есть *простой* корень уравнения (12.32).

Рассмотрим теперь какой-либо другой корень $x \neq p$, $(x \neq 0)$. Тогда, полагая $|x| = q > 0$, имеем

$$1 = \frac{b_{k_1}}{x^{k_1}} + \frac{b_{k_2}}{x^{k_2}} + \ldots + \frac{b_{k_m}}{x^{k_m}} \leqslant \frac{b_{k_1}}{q^{k_1}} + \frac{b_{k_2}}{q^{k_2}} + \ldots + \frac{b_{k_m}}{q^{k_m}}, \quad (12.35)$$

откуда видно, что $F(q) \geqslant 0$. Остается доказать, что $F(q) > 0$, т. е. $F(q) > F(p)$ и, значит, $q < p$. Допустим, от противного, что $F(q) = 0$. Тогда в формуле (12.35) должен стоять знак равенства, и поэтому все дроби $\frac{b_{k_\mu}}{x^{k_\mu}}$ должны быть положительны. По формуле (12.33) имеем

$$\left(\frac{b_{k_1}}{x^{k_1}}\right)^{s_1} \left(\frac{b_{k_2}}{x^{k_2}}\right)^{s_2} \ldots \left(\frac{b_{k_m}}{x^{k_m}}\right)^{s_m} = \frac{b_{k_1}^{s_1} b_{k_2}^{s_2} \ldots b_{k_m}^{s_m}}{x},$$

откуда следует, что вместе с рассматриваемым произведением должен быть положительным также и x. Но это означает, что $x = p$, вопреки условию. Теорема 12.2 доказана [1]).

---

[1]) Условия теоремы 12.2 относительно индексов b_ν, отличных от нуля, очевидно, существенны. Действительно, если в левой части уравнения (12.32) стоит многочлен от x^k, $k > 1$, то это уравнение имеет k корней, по модулю равных p.

ГЛАВА 13

n РАЗЛИЧНЫХ УЗЛОВ ИНТЕРПОЛЯЦИИ

Оценки погрешностей

1. Пусть функция $w = f(x)$ определена на J_x; пусть $f(\zeta) = 0$, $\zeta \in J_x$ и $f'(x) \neq 0$ при всех $x \in J_x$. Допустим, что на J_x заданы n различных узлов интерполяции x_ν ($\nu = 1, \ldots, n$), лежащих в некоторой достаточно малой окрестности точки ζ; обозначим $f(x_\nu) = y_\nu$ ($\nu = 1, \ldots, n$). Функцию $x = \varphi(w)$, обратную к функции $w = f(x)$, будем аппроксимировать интерполяционным многочленом Лагранжа степени $n-1$, $T_{n-1}(w) = T(w)$, заданным условиями $T(y_\nu) = x_\nu$ ($\nu = 1, \ldots, n$). Применяя формулу (1.7) с заменой x_ν на y_ν, $f(x_\nu)$ на x_ν, x на 0, $F(x)$ на $\Phi(w) = \prod\limits_{\nu=1}^{n}(w - y_\nu)$ и полагая $T(0) = x_{n+1}$, получим

$$x_{n+1} = T(0) = (-1)^{n-1} \prod_{\nu=1}^{n} y_\nu \sum_{\mu=1}^{n} \frac{x_\mu}{y_\mu} \frac{1}{\Phi'(y_\mu)}. \qquad (13.1)$$

Предположим теперь, что $f^{(n)}(x)$ непрерывна на J_x. Погрешность приближения x_{n+1} к корню $\zeta = \varphi(0)$ мы найдем из (2.2):

$$\zeta - x_{n+1} = (-1)^n S \prod_{\nu=1}^{n} y_\nu, \qquad (13.2)$$

где

$$S = \frac{1}{n!}\varphi^{(n)}(\eta), \quad \eta \in (0, y_1, \ldots, y_n). \qquad (13.3)$$

Так как

$$y_\nu = f(x_\nu) = (x_\nu - \zeta) f'(\xi_\nu), \quad \xi_\nu \in (x_\nu, \zeta),$$

то формулу (13.2) можно записать в виде

$$\zeta - x_{n+1} = S \prod_{\nu=1}^{n} (\zeta - x_\nu) \prod_{\nu=1}^{n} f'(\xi_\nu),$$

$$\frac{\zeta - x_{n+1}}{\prod\limits_{\nu=1}^{n}(\zeta - x_\nu)} = S \prod_{\nu=1}^{n} f'(\xi_\nu). \tag{13.4}$$

Отсюда следует, в частности, что если все $x_\nu \to \zeta$ $(\nu = 1, \ldots, n)$, то

$$\frac{\zeta - x_{n+1}}{\prod\limits_{\nu=1}^{n}(\zeta - x_\nu)} \to \frac{1}{n!}\varphi^{(n)}(0) f'(\zeta)^n \quad (x_\nu \to \zeta, \quad \nu = 1, \ldots, n). \tag{13.5}$$

2. Пусть k есть верхняя грань модуля правой части (13.4). Тогда

$$|\zeta - x_{n+1}| \leqslant k \prod_{\nu=1}^{n} |\zeta - x_\nu|.$$

Это неравенство можно записать в виде

$$K|\zeta - x_{n+1}| \leqslant \prod_{\nu=1}^{n} (K|\zeta - x_\nu|), \tag{13.6}$$

где $K = k^{\frac{1}{n-1}}$. Если все x_ν $(\nu = 1, \ldots, n)$ выбраны достаточно близко к ζ, то все множители в правой части (13.6) будут меньше 1. Введем ε_ν по формуле

$$K|\zeta - x_\nu| = u^{\varepsilon_\nu} \quad (\nu \geqslant 1), \tag{13.7}$$
$$u = \max(K|\zeta - x_1|, K|\zeta - x_2|).$$

Тогда из (13.6) и (13.7) имеем

$$u^{\varepsilon_{n+1}} \leqslant u^{\varepsilon_1 + \ldots + \varepsilon_n}. \tag{13.8}$$

Итерации с n различными узлами интерполяции

3. В дальнейшем мы предполагаем, что $u < 1$. Тогда

$$\varepsilon_{n+1} \geqslant \varepsilon_1 + \ldots + \varepsilon_n. \tag{13.9}$$

Принципиально возможны два различных способа построения дальнейших приближений к корню. Мы можем начать с $n = 2$ и продолжать с $n = 3, 4, \ldots$, добавляя по одной точке на каждом шаге. Пусть

$$s_n = \sum_{\nu=1}^{n} \varepsilon_\nu \quad (n \geqslant 2).$$

Тогда из неравенства (13.9) следует, что

$$s_{n+1} = s_n + \varepsilon_{n+1} \geqslant 2s_n.$$

При этом $s_2 = \varepsilon_1 + \varepsilon_2$ и $\min(\varepsilon_1, \varepsilon_2) = 1$. Следовательно,

$$s_n \geqslant 2^{n-2} s_2 \geqslant 2^{n-1},$$

и на каждом шаге имеем

$$K|\zeta - x_{n+1}| \leqslant u^{2^{n-1}}. \tag{13.10}$$

Наша погрешность убывает по квадратичному закону. Каждый шаг осуществляется посредством одного горнера, т. е. путем вычисления f_{n+1}; поэтому индекс эффективности такого метода равен 2. Однако этот метод нельзя рекомендовать для расчетов, так как он не допускает никакого контроля, кроме контроля путем повторения всех вычислений. Альтернативой такого контроля мог бы быть контроль по формуле Ньютона, но он требовал бы лишнего горнера [1]).

4. Другой возможный способ состоит в использовании фиксированного числа n узлов на каждом шаге: по точкам $x_1, \ldots, x_n$ находят x_{n+1}, затем по точкам $x_2, \ldots, x_{n+1}$ находят x_{n+2} и т. д. Как и в предыдущем случае, мы затрачиваем на каждом шаге один горнер.

Если, начав с $x_i, x_{i+1}, \ldots, x_{i+n-1}$, мы вычисляем x_{i+n}, то в обозначениях (13.7) имеем

$$u^{\varepsilon_{i+n}} \leqslant u^{\varepsilon_i + \varepsilon_{i+1} + \cdots + \varepsilon_{i+n-1}}.$$

И так как, по предположению, $u < 1$, то

$$\varepsilon_{i+n} \geqslant \varepsilon_i + \varepsilon_{i+1} + \ldots + \varepsilon_{i+n-1}. \tag{13.11}$$

5. Так как все $\varepsilon_\nu > 0$, то ε_ν монотонно возрастают вместе с ν. Если бы ε_ν имели предел A при $\nu \to \infty$, то из (13.11) следовало бы, что $A \geqslant nA$; поэтому

$$\varepsilon_\nu \to \infty \quad \text{при} \quad \nu \to \infty$$

и

$$x_\nu - \zeta \to 0 \qquad (\nu \to \infty). \tag{13.12}$$

---

[1]) На первый взгляд кажется, что описанный метод страдает и другим недостатком, так как, по-видимому, требует вычисления первых y_ν с произвольно большой точностью. Однако, как следует из приложения VIII, этот недостаток не является существенным.

6. Положим

$$\ln \frac{1}{|\zeta - x_\nu|} = y_\nu. \tag{13.13}$$

Применив формулу (13.4) к последовательности x_i, x_{i+1}, ..., x_{i+n-1}, x_{i+n} и взяв логарифмы модулей обеих частей (13.4), получим

$$y_{i+n} = y_i + y_{i+1} + \ldots + y_{i+n-1} + k_i, \tag{13.14}$$

где k_i при $\varphi^{(n)}(0) \neq 0$ имеют конечный предел

$$\beta = -\ln\left[\frac{1}{n!}|\varphi^{(n)}(0)|\right] - n \ln |f'(\zeta)|.$$

7. Формула (13.14) дает линейное разностное уравнение типа (12.1), где k_i образуют ограниченную последовательность.

Покажем теперь, что все условия теоремы 12.1 выполнены. Характеристическим многочленом уравнения (13.14) является

$$f_n(x) \equiv x^n - x^{n-1} - \ldots - x - 1. \tag{13.15}$$

Исследуем корни этого многочлена при $n > 1$.

Исследование корней характеристического многочлена

8. По правилу Декарта многочлен (13.15) имеет точно один положительный корень μ_n. Так как $f_n(1) = -(n-1)$, $f_n(2) = 1$, то

$$1 < \mu_n < 2. \tag{13.16}$$

Из формулы (13.15) следует, что $f_n(x) = xf_{n-1}(x) - 1$, поэтому $f_n(\mu_{n-1}) = -1$. Итак,

$$1 < \mu_{n-1} < \mu_n < 2 \qquad (n = 3, 4, \ldots). \tag{13.17}$$

Для дальнейшего исследования μ_n мы воспользуемся неравенством

$$\frac{2^{n+1}}{n+1} > \left(1 + \frac{1}{n}\right)^n \qquad (n = 2, 3, \ldots). \tag{13.18}$$

Справедливость этого неравенства при $n = 2$ проверяется непосредственно. При $n \geqslant 3$ справедливость неравенства (13.18) вытекает из того, что его левая часть $\geqslant 4$, а правая $< e < 4$.

9. При $n > 1$ из (13.15) имеем

$$F_n(x) \equiv (x-1) f_n(x) = x^{n+1} - 2x^n + 1 = \\ = 1 - (2-x)x^n. \qquad (13.19)$$

Полагая $x = \dfrac{2n}{n+1}$ и применяя неравенство (13.18), получаем

$$F_n\left(\frac{2n}{n+1}\right) = 1 - \frac{2^{n+1}}{n+1}\left(\frac{n}{n+1}\right)^n = \\ = 1 - \frac{2^{n+1}}{n+1}\left(1 + \frac{1}{n}\right)^{-n} < 0. \qquad (13.19a)$$

Отсюда следует, что $f_n\left(\dfrac{2n}{n+1}\right) < 0$ и, значит,

$$\frac{2n}{n+1} < \mu_n < 2; \quad \mu_n \uparrow 2 \quad (n \to \infty). \qquad (13.20)$$

С другой стороны,

$$F_n'(x) = (n+1)\, x^{n-1}\left(x - \frac{2n}{n+1}\right).$$

Сравнивая это равенство с неравенством (13.19а), видим, что $F_n(x)$, а значит, и $f_n(x)$ не имеют кратных корней.

10. Докажем теперь лемму о многочлене (13.15).

Лемма 13.1. Если q_n — максимум модулей всех корней многочлена (13.15), отличных от μ_n, то при $n > 2$

$$q_n < \mu_n - 1 < 1 \quad (n > 2). \qquad (13.21)$$

Замечание. При $n=2$ из главы 3, п. 10 следует, что

$$q_2 = |t_2| = t_1 - 1 = \mu_2 - 1.$$

Доказательство леммы. Положив, для простоты, $\mu_n = \mu$, рассмотрим многочлен

$$\frac{f_n(x)}{x-\mu} = c_0 x^{n-1} + \ldots + c_\nu x^{n-\nu-1} + \ldots + c_{n-1}, \qquad (13.22)$$

где $c_0 = 1$;

$$c_\nu = \mu^\nu - \mu^{\nu-1} - \ldots - \mu - 1 \quad (1 \leqslant \nu \leqslant n-1). \qquad (13.23)$$

Так как

$$f_n(\mu) = \mu^n - \mu^{n-1} - \ldots - \mu - 1 = 0,$$

то, разделив это равенство на $\mu^{n-\nu}$ и перенеся последние $n-\nu$ членов в правую часть, получим

$$c_\nu = \frac{1}{\mu} + \frac{1}{\mu^2} + \ldots + \frac{1}{\mu^{n-\nu}} \quad (\nu = 0,\ 1,\ \ldots,\ n-1).$$

Мы видим, что все коэффициенты c_ν положительны и что

$$c_{\nu-1} = c_\nu + \frac{1}{\mu^{n-\nu+1}} > c_\nu \quad (\nu = 1,\ \ldots,\ n-1).$$

11. Из формулы (13.23) вытекает, что

$$c_\nu = \mu c_{\nu-1} - 1;$$

$$\frac{c_\nu}{c_{\nu-1}} = \mu - \frac{1}{c_{\nu-1}}.$$

Сравнивая это отношение с отношением

$$\frac{c_{\nu+1}}{c_\nu} = \mu - \frac{1}{c_\nu},$$

находим, что

$$\frac{c_{\nu+1}}{c_\nu} < \frac{c_\nu}{c_{\nu-1}} \qquad (\nu = 1,\ \ldots,\ n-2). \tag{13.24}$$

12. Воспользуемся теперь следующей теоремой, принадлежащей (за исключением второй части) Энестрему и Какейа:

Если в уравнении

$$g(x) \equiv a_0 x^{n-1} + \ldots + a_{n-1} = 0 \tag{13.25}$$

все коэффициенты a_ν положительны, то для каждого корня ξ этого уравнения справедлива оценка

$$|\xi| \leqslant \gamma = \max_{1 \leqslant \nu \leqslant n-1} \frac{a_\nu}{a_{\nu-1}}. \tag{13.26}$$

Пусть $k_1,\ k_2,\ \ldots,\ k_m$ — те индексы k, для которых $\frac{a_k}{a_{k-1}} < \gamma$. Тогда, если наибольший общий делитель чисел $n,\ k_1, \ldots, k_m$ равен 1, то в (13.26) *имеет место строгое неравенство* [1]).

---

[1]) Пример многочлена $x^3 + \gamma x^2 + cx + \gamma c = (x+\gamma)(x^2+c)$, где $0 < c < \gamma^2$, показывает, что без дополнительных условий нельзя требовать строгого неравенства в (13.26).

Применяя эту теорему к многочлену (13.22) и учитывая неравенство (13.24), а также то, что наибольший общий делитель чисел 2, 3, ..., n ($n > 2$) равен 1, получаем

$$q_n < \max\left(\frac{c_1}{c_0}, \frac{c_2}{c_1}, \ldots, \frac{c_{n-1}}{c_{n-2}}\right) = \frac{c_1}{c_0},$$

т. е.

$$q_n < c_1 = \mu_n - 1, \quad \text{ч. т. д.}$$

13. Для полноты изложения мы даем здесь доказательство приведенной выше теоремы, так как вопрос о строгом неравенстве в (13.26) в литературе не встречается.

Рассмотрим многочлен

$$(x - \gamma)\, g(x) = a_0 x^n - (\gamma a_0 - a_1)\, x^{n-1} - \ldots$$
$$\ldots - (\gamma a_{\nu-1} - a_\nu)\, x^{n-\nu} - \ldots - a_{n-1}\gamma.$$

Так как по определению числа γ все выражения $\gamma a_{\nu-1} - a_\nu$ ($\nu = 1, \ldots, n-1$) неотрицательны, то γ является простым и единственным положительным корнем многочлена $(x - \gamma)\, g(x)$ и неравенство (13.26) следует из теоремы 12.2. Если выполнено дополнительное условие второй части теоремы, то индексы отличных от нуля коэффициентов многочлена $(x - \gamma)\, g(x)$ имеют наибольшим общим делителем 1. Поэтому единственный положительный корень γ этого многочлена будет *больше* модулей всех остальных его корней, т. е. больше модулей *всех корней* многочлена (13.25).

14. Мы можем теперь применить к решениям y_ν уравнения (13.14) теорему 12.1 при $u_1 = \mu_n$ и $s = 1$. По формуле (12.13), при $\nu \to \infty$

$$\frac{y_\nu}{\mu_n^\nu} \to \alpha > 0$$

и, следовательно,

$$\ln|\zeta - x_\nu| \sim -\alpha\mu_n^\nu, \tag{13.27}$$

$$\frac{\ln|\xi - x_{\nu+1}|}{\ln|\zeta - x_\nu|} \to \mu_n. \tag{13.28}$$

Итак, если мы применяем n-точечную интерполяцию, то на каждом шаге погрешность возводится (асимптотически) в степень μ_n, так что *индекс эффективности нашего итерационного процесса равен* μ_n.

15. Для получения более точной [вида (12.16)] оценки решений y_ν рассмотрим подробнее свободный член k_i уравнения (13.14). Предположим, что

$$\varphi^{(n)}(0) \neq 0, \quad f'(\zeta) \neq 0. \tag{13.29}$$

Применяя (13.4) к последовательности $x_i, \ldots, x_{i+n}$, получаем

$$\frac{|\zeta - x_i| \ldots |\zeta - x_{i+n-1}|}{|\zeta - x_{i+n}|} =$$

$$= \frac{n!}{|\varphi^{(n)}(\eta^{(i)})|} \prod_{\nu=1}^{n} \frac{1}{|f'(\xi_\nu^{(i)})|},$$

откуда

$$k_i = y_{i+n} - y_i - \ldots - y_{i+n-1} =$$
$$= \ln\left|n!\,\frac{f'(\zeta)^{-n}}{\varphi^{(n)}(0)}\right| - \ln\left|\frac{\varphi^{(n)}(\eta^{(i)})}{\varphi^{(n)}(0)}\right| -$$
$$- \sum_{\nu=1}^{n} \ln\left|\frac{f'(\xi_\nu^{(i)})}{f'(\zeta)}\right|. \tag{13.30}$$

16. Все $\xi_\nu^{(i)}$ лежат между наибольшим и наименьшим из $(n+2)$ чисел

$$x_i, \ldots, x_{i+n}, \zeta$$

и стремятся к ζ при $i \to \infty$, если $x_\nu \to \zeta$ $(\nu \to \infty)$. Из формулы (13.3) видно, что

$$|\eta^{(i)}| \leqslant \max(|f(x_i)|, \ldots, |f(x_{i+n-1})|).$$

Поэтому

$$\xi_\nu^{(i)} - \zeta = O(|x_i - \zeta|), \quad \eta^{(i)} = O(|x_i - \zeta|),$$

так как, начиная с некоторого i величина $|x_i - \zeta|$ убывает.

Предполагая, что $\varphi^{(n+1)}$ непрерывна в окрестности нуля, т. е. что $f^{(n+1)}$ непрерывна в окрестности корня ζ, мы имеем

$$\frac{\varphi^{(n)}(\eta^{(i)})}{\varphi^{(n)}(0)} = 1 + O(|x_i - \zeta|),$$

$$\frac{f'(\xi_\nu^{(i)})}{f'(\zeta)} = 1 + O(|x_i - \zeta|),$$

$$-\ln\left|\frac{\varphi^{(n)}(\eta^{(i)})}{\varphi^{(n)}(0)}\right| - \sum_{\nu=1}^{n} \ln\left|\frac{f'(\xi_\nu^{(i)})}{f'(\zeta)}\right| = O(|x_i - \zeta|),$$

и, значит,

$$k_i - \beta = k_i - \ln\left|\frac{n!}{\varphi^{(n)}(0)\, f'(\zeta)^n}\right| = O(|x_i - \zeta|). \quad (13.31)$$

Положив теперь в уравнении (13.30)

$$y_i = v_i + \frac{\beta}{1-n} = v_i + \beta_1,$$

мы получаем уравнение

$$v_{i+n} - v_i - \ldots - v_{i+n-1} = k_i - \beta, \quad (13.32)$$

в котором правая часть есть $O(s^i)$ при любом сколь угодно малом положительном s.

17. Так как многочлен (13.15) является характеристическим многочленом также и для разностного уравнения (13.32), то мы видим, что выполнены все условия теоремы 12.1 с произвольно малым s. Из формулы (12.16) при $u_1 = \mu_n$ и $m = 1$ имеем

$$v_\nu = \alpha\mu_n^\nu + O(q_n^\nu),$$

$$y_\nu = \beta_1 + \alpha\mu_n^\nu + O(q_n^\nu),$$

$$|x_\nu - \zeta| = e^{-\beta_1} e^{-\alpha\mu_n^\nu}\left[1 + O(q_n^\nu)\right], \quad (13.33)$$

$$\frac{|x_{\nu+1} - \zeta|}{|x_\nu - \zeta|^{\mu_n}} = e^{(\mu_n - 1)\beta_1} + O(q_n^\nu), \quad (13.34)$$

где

$$0 < q_n \leqslant \mu_n - 1 < 1. \quad (13.35)$$

18. Следующая таблица [1]) чисел μ_n показывает, что обычно не сто́ит выбирать $n > 3$ или $n > 4$:

n	μ_n	n	μ_n
2	1,61803	9	1,99803
3	1,83929	10	1,99902
4	1,92756	11	1,99952
5	1,96595	12	1,99976
6	1,98358	13	1,99988
7	1,99196	14	1,99994
8	1,99603	15	1,99997

[1]) Вычисления проведены Б. Х. Уолтер в Лаборатории вычислительной математики Национального бюро стандартов, Вашингтон

ГЛАВА 14

$n+1$ СОВПАДАЮЩИЙ УЗЕЛ ИНТЕРПОЛЯЦИИ И ТЕЙЛОРОВСКОЕ РАЗЛОЖЕНИЕ КОРНЯ

Постановка задачи

1. Пусть функция $w = f(z)$ определена на J_z. Предположим, что заданы $n+1$ совпадающий узел интерполяции, т. е. точка z_0 и величины

$$f(z_0) = w_0, \quad f'(z_0), \quad f''(z_0), \quad \ldots, \quad f^{(n)}(z_0).$$

Разложим z в окрестности точки w_0 в ряд Тейлора:

$$z = \varphi(w) = z_0 + \sum_{\nu=1}^{\infty} \frac{(w-w_0)^\nu}{\nu!} \varphi^{(\nu)}(w_0). \tag{14.1}$$

При $w = 0$ получаем

$$\zeta - z_0 = \sum_{\nu=1}^{\infty} \frac{(-1)^\nu}{\nu!} w_0^\nu \varphi^{(\nu)}(w_0). \tag{14.2}$$

Частичная сумма этого ряда с n членами дает аппроксимацию корня, использующую только данные величины. При этом оценка погрешности может быть получена из формулы (2.2), если положить в ней $y_1 = y_2 = \ldots = y_n = w_0$. Выражение $\varphi^{(n+1)}$ через $f^{(k)}$ $(k = 0, 1, \ldots, n+1)$ можно получить по формуле (2.5) и по таблице п. 8 гл. 2 или непосредственно по формуле (III. 5) приложения III. Однако с ростом n вычисление по этим формулам становится очень сложным и, кроме того, вопрос о существовании корня требует отдельного исследования.

Одна теорема об обратной функции и конформном отображении

2. С другой стороны, применение ряда (14.2) в практических вычислениях может оказаться удобным, если

вычисление производных в точке z_0 проще, чем при других значениях z. Конечно, все используемые величины $f^{(k)}(z_0)$ должны быть при этом вычислены сразу с достаточно большим числом знаков.

Мы исследуем вопросы сходимости ряда (14.2) и существования корня ζ, а также получим практически удобные оценки остаточного члена, пользуясь теорией функций комплексного переменного.

Докажем сначала следующую теорему.

Теорема 14.1. Пусть функция $w = f(z)$ аналитична в замкнутом круге K_z: $|z - z_0| \leqslant r$. Пусть $f(z_0) = w_0$. Если

$$M = M(r) = \max_{K_z} |f''(z)| \leqslant \frac{2|f'(z_0)|}{r}, \tag{14.3}$$

то ряд для обратной функции

$$z = \varphi(w) = z_0 + \sum_{\nu=1}^{\infty} \frac{(w - w_0)^\nu}{\nu!} \varphi^{(\nu)}(w_0) \tag{14.4}$$

сходится в круге K_w: $|w - w_0| < R$, где

$$R = R(r) = |f'(z_0)|\, r - \frac{Mr^2}{2}, \tag{14.5}$$

и $\varphi(w)$ удовлетворяет в этом круге неравенству

$$|\varphi(w) - z_0| < r. \tag{14.6}$$

3. Доказательство. Без ограничения общности можно считать, что $z_0 = w_0 = 0$. Мы будем писать f_0' вместо $f'(0)$. Положим

$$f(z) = f_0' z + T(z). \tag{14.7}$$

Здесь $T(z)$, как остаточный член формулы Тейлора, может быть представлен в виде

$$T(z) = \int_0^z (z - u) f''(u)\, du. \tag{14.8}$$

Действительно, интегрирование по частям дает

$$\int_0^z (z-u) f''(u)\,du = (z-u) f'(u) \Big|_0^z + \int_0^z f'(u)\,du =$$
$$= f(z) - f_0' z.$$

Подстановка $u = tz$ приводит интеграл (14.8) к виду

$$T(z) = z^2 \int_0^1 (1-t) f''(tz)\,dt. \tag{14.9}$$

4. Для любой точки z, лежащей на границе K_z, справедливо неравенство

$$|T(z)| \leqslant r^2 M \int_0^1 (1-t)\,dt = \frac{Mr^2}{2}. \tag{14.10}$$

С другой стороны, для линейной функции

$$L(z) = f_0' z - w \tag{14.11}$$

при $|w| < R$ и при $|z| = r$ имеет место неравенство

$$|L(z)| \geqslant |f_0'|\, r - |w| > \frac{Mr^2}{2},$$

вытекающее из (14.5). Поэтому, если z лежит на границе K_z, а w внутри K_w, то

$$|L(z)| > |T(z)| \quad (|w| < R). \tag{14.12}$$

5. Воспользуемся теперь известной теоремой Руше. *Пусть две функции $g(z)$ и $h(z)$ аналитичны в односвязной области G и пусть простая замкнутая кривая C лежит внутри G. Если $|g(z)| > |h(z)|$ вдоль всей кривой C, то функции $g(z)$ и $g(z)+h(z)$ имеют одинаковое число нулей в подобласти G, ограниченной кривой C.*

Применим эту теорему к функциям $g(z) = L(z)$ и $h(z) = T(z)$, взяв в качестве кривой C окружность $|z| = r$. Из формулы (14.11) видно, что единственным нулем функции $L(z)$ является

$$z_* = \frac{w}{f_0'},$$

причем z_* лежит внутри K_z, так как при $|w|<R$ по (14.5)

$$|z_*| = \frac{|w|}{|f_0'|} < r - \frac{Mr^2}{2|f_0'|} < r.$$

Поэтому внутри K_z лежит точно один нуль функции

$$L(z) + T(z) = [f_0' z - w] + [f(z) - f_0' z] = f(z) - w,$$

т. е. точно один корень уравнения $f(z) = w$.

Таким образом, функция $w = f(z)$ определяет конформное отображение некоторой подобласти круга K_z, содержащей 0, на всю внутренность круга K_w. Мы видим, что обратная функция $z = \varphi(w)$ аналитична в круге K_w и удовлетворяет в нем неравенству (14.6). Теорема 14.1 доказана.

6. В приложениях теоремы 14.1 может оказаться желательным такой выбор r, чтобы R было возможно бо́льшим. Хотя это лежит в стороне от нашего исследования, скажем несколько слов по этому вопросу. Хорошо известно, что функция

$$M(r) = \max_{|z| \leqslant r} |f''(z)|$$

непрерывна и строго возрастает, если только $f''(z)$ не постоянна.

Если теперь при выбранном значении r

$$M(r) > \frac{|f'(z_0)|}{r}, \tag{14.13}$$

то оценка (14.5) может быть улучшена. Действительно, в этом случае существует такое положительное число $\varrho < r$, что $\varrho M(\varrho) = |f'(z_0)|$. При этом условия теоремы 14.1 выполнены для ϱ, и мы находим для $R(\varrho)$ значение

$$R(\varrho) = |f'(z_0)|\varrho - \frac{1}{2} M(\varrho)\varrho^2 =$$

$$= \frac{1}{2}|f'(z_0)|\varrho = \frac{|f'(z_0)|^2}{2M(\varrho)}. \tag{14.14}$$

Эта величина больше, чем $R_* = \frac{|f'(z_0)|^2}{2M(r)}$, а значит, заведомо больше, чем $R(r)$, так как

$$R_* - R(r) = \frac{|f'(z_0)|^2}{2M(r)} - |f'(z_0)|r + \frac{M(r)r^2}{2} =$$

$$= \frac{[|f'(z_0)| - rM(r)]^2}{2M(r)} > 0.$$

На практике вместо $R(r)$ берут величину $R_* = \frac{|f'(z_0)|^2}{2M(r)}$, если только r удовлетворяет условию (14.13).

Теорема о погрешностях тейлоровской аппроксимации корня

7. Следующая теорема, вытекающая из теоремы 14.1, дает хорошее практическое решение проблем, связанных с применением ряда (14.2) для приближенного вычисления корня.

Теорема 14.2. Пусть $f(z)$ — нелинейная функция, аналитическая в z_0, $f(z_0) = w_0$, $|f'(z_0)| = f'_0$. Обозначим через σ радиус сходимости тейлоровского разложения функции $f(z)$ в окрестности z_0. Положим при $0 < r < \sigma$

$$M(r) = \max_{|z-z_0| \leqslant r} |f''(z)| \tag{14.15}$$

и

$$R(r) = f'_0 r - \frac{1}{2} M(r) r^2. \tag{14.16}$$

Если при некотором $r < \sigma$ имеет место неравенство $R(r) > |w_0|$, то существует один и только один корень ζ уравнения $f(z)=0$, удовлетворяющий неравенству $|\zeta - z_0| < r$. При этом ряд (14.2) сходится к величине $\zeta - z_0$ и мажорируется рядом

$$r \sum_{\nu=1}^{\infty} \frac{|w_0|^\nu}{R(r)^\nu}.$$

В частности, погрешность приближения

$$E_n = \zeta - \left[z_0 + \sum_{\nu=1}^{n} \frac{(-1)^\nu}{\nu!} w_0^\nu \varphi^{(\nu)}(w_0) \right] \tag{14.17}$$

оценивается величиной

$$|E_n| \leqslant \left[\frac{|w_0|}{R(r)}\right]^{n+1} \frac{r}{1-\frac{|w_0|}{R(r)}}. \tag{14.18}$$

8. Доказательство. Так как по теореме 14.1 радиус сходимости ряда (14.1) не меньше $R(r)$, то ясно, что ряд (14.2) сходится.

По неравенству Коши получаем из формулы (14.6)

$$\left|\frac{1}{\nu!}\varphi^{(\nu)}(w_0)\right| < \frac{r}{R(r)^\nu} \quad (\nu = 1, 2, \ldots). \tag{14.19}$$

Отсюда следует, что ряд $r\sum\limits_{\nu=1}^{\infty}\frac{|w_0|^\nu}{R(r)^\nu}$ является мажорантой для ряда (14.2) и, в частности,

$$\left|\zeta - z_0 - \sum_{\nu=1}^{n}\frac{(-1)^\nu}{\nu!}w_0^\nu\varphi^{(\nu)}(w_0)\right| \leqslant r\sum_{\nu=n+1}^{\infty}\left[\frac{|w_0|}{R(r)}\right]^\nu =$$

$$= \left[\frac{|w_0|}{R(r)}\right]^{n+1}\frac{r}{1-\frac{|w_0|}{R(r)}}, \text{ ч. т. д.}$$

Исследование условий теоремы 14.2

9. Исследуем теперь условие $|w_0| < R(r)$. Разделив это неравенство на $\frac{f_0'}{2}$ и применив обозначения

$$h_0 = \frac{f(z_0)}{f'(z_0)} = \frac{w_0}{f'(z_0)}, \quad |h_0| = \frac{|w_0|}{f_0'} = \tau, \tag{14.20}$$

получим условие

$$2\tau < r\left[2 - \frac{rM(r)}{f_0'}\right]; \tag{14.21}$$

отсюда непосредственно следует, что $r > \tau$. Разрешая неравенство (14.21) относительно $rM(r)$, получаем

$$rM(r) < 2\left(1 - \frac{\tau}{r}\right)f_0', \tag{14.22}$$

$$M(r) < 2\left(\frac{1}{r} - \frac{\tau}{r^2}\right)f_0'. \tag{14.23}$$

Предположим теперь, что (14.23) выполнено при некотором $r > 2\tau$. Если заменить это r на 2τ, то левая часть выражения (14.23) уменьшится, а правая увеличится, так как правая часть монотонно убывает с ростом r при $r > 2\tau$.

Следовательно, в этом случае замена r на 2τ *усиливает* неравенство (14.22). Таким образом, мы приходим к очень удобному достаточному условию сходимости разложения (14.2)

$$r = 2\tau < \sigma; \quad 2\tau M(2\tau) < f_0'. \tag{14.24}$$

Сравним теперь условия сходимости (14.22) и (14.24) с условием теоремы 7.1:

$$2\tau M \leqslant f_0', \tag{14.25}$$

где τ задается формулой (14.20), а $M = \max|f''(z)|$ в круге $K_0 : |z - z_1| = |z - z_0 - h_0| \leqslant \tau$. Так как $M(2\tau)$ есть $\max|f''(z)|$ в круге $K^{(2\tau)} : |z - z_0| \leqslant 2\tau$, который содержит круг K_0 (см. рис. 6), то неравенство (14.24) не следует из (14.25).

Из доказательства теоремы 7.1 следует, однако, что условие (14.24) будет выполнено, если только мы заменим z_0 на следующее ньютоновское приближение z_1 (исключением является лишь случай $f(z) = a(z - \zeta)^2$).

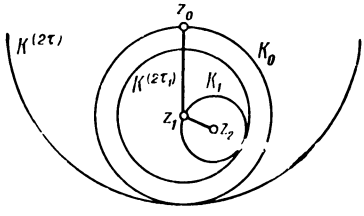

Рис. 6.

Действительно, из неравенства (7.10) следует, что круг $K^{2\tau_1}$ радиуса $2\tau_1 = 2|h_1|$ с центром в z_1 лежит в круге K_0 (см. рис. 6). Поэтому число M в (14.25) не меньше числа $M(2\tau_1)$, вычисленного в круге $K^{2\tau_1}$ с центром z_1.

Заметим, что из неравенства (7.9) следует лишь смешанное неравенство:

$$2\tau_1 M(2\tau_1) \leqslant f_0'.$$

Но при выводе неравенства (7.9) мы опирались на неравенства (7.5) и (7.7), в которых знак равенства не имеет места, если только $f''(z)$ не постоянна, т. е. если $f(z)$ не является квадратичным многочленом. Очевидно, что все сказанное выше справедливо для любого z_ν, $\nu > 1$. Если мы используем теперь последнее замечание к приложению VI, то получим

Следствие. Если выполнены условия теоремы 7.1, то, заменив в тейлоровском разложении (14.2) z_0 и $f(z_0)$ на z_ν и $f(z_\nu)$, мы получим при любом $\nu > 0$ сходящееся разложение. Исключением является только случай, когда $f(z)$ есть квадратичный многочлен с двойным корнем.

При получении приближений с помощью указанного разложения не всегда удобно брать $r = 2|h_\nu|$, так как при этом по (14.18) получается весьма завышенная оценка погрешности. Обычно в качестве r можно взять расстояние от z_ν до окружности $|z - z_1| = |h_0|$ и получить таким способом весьма удовлетворительную оценку, если ν достаточно велико.

10. Оценка (14.18) обычно бывает намного больше, чем действительная погрешность, хотя эта оценка имеет тот же порядок малости, что и погрешность. С другой стороны, вычислитель обычно считает — и очень часто правильно,— что действительная ошибка имеет порядок следующего члена ряда. Мы не можем сделать такой вывод прямо из формулы (14.18), так как неравенство (14.19) дает обычно очень грубую оценку. Однако можно продолжать рассуждение следующим образом.

Применим (14.18) при некотором $m > n$; мы получим

$$|E_n| \leqslant \left[\frac{|w_0|}{R(r)}\right]^{m+1} \frac{r}{1 - \frac{|w_0|}{R(r)}} + \sum_{\nu=n+1}^{m} \frac{|w_0|^\nu}{\nu!} |\varphi^{(\nu)}(w_0)|. \quad (14.26)$$

При употреблении этой формулы, конечно, нет необходимости вычислять величины членов, стоящих под знаком суммы в правой части, более чем с одной или двумя значащими цифрами. Можно применять даже грубые оценки, если они лучше, чем (14.19).

11. Пример. Уравнение

$$f(x) \equiv x^3 - 2x - 5 = 0$$

имеет точно один положительный корень:

$$\zeta = 2{,}0945514815423 \ldots .$$

Здесь $f'(x) = 3x^2 - 2$; $f''(x) = 6x$. Возьмем $z_0 = 2$, тогда

$$z_0 = 2; \quad w_0 = -1; \quad f'_0 = 10, \quad f''(z_0) = 12.$$

Очевидно, что $M(r) = 12 + 6r$. Если взять $r = \frac{1}{2}$, то $R(r) = 3{,}125$.

Первые три производные функции $\varphi(w)$ в точке w_0 равны

$$\varphi'(w_0) = 0{,}1; \quad \varphi''(w_0) = -0{,}012; \quad \varphi'''(w_0) = 0{,}00372.$$

При $n = 3$ имеем

$$\zeta \sqsupset 2 + 0{,}1 - 0{,}006 + 0{,}00062 = 2{,}09462.$$

Погрешность здесь равна 0,00007, в то время как формула (14.18) дает оценку 0,00772.

Если положим

$$\begin{gathered} z_0 = 2{,}1; \quad w_0 = 0{,}061; \quad f'_0 = 11{,}23; \\ f''(z_0) = 12{,}6; \quad r = \frac{1}{2}, \end{gathered} \tag{14.27}$$

то $R = 3$ удовлетворяет условиям теоремы 14.1.

Мы получаем

$$\begin{gathered} \varphi'_0 = 0{,}089047195; \quad \varphi''_0 = -0{,}00889674773; \\ \varphi'''_0 = 0{,}002289382342; \quad \zeta \sqsupset 2{,}094551482097. \end{gathered} \tag{14.28}$$

Погрешность здесь равна $5{,}5 \cdot 10^{-10}$, а оценка по формуле (14.18) дает $9{,}33 \cdot 10^{-8}$.

Если в случае (14.27) взять еще два члена ряда, то получим значение корня с 15 верными знаками после запятой.

Другой метод аппроксимации корней уравнения $f(z) = 0$ будет исследован в приложении X.

ГЛАВА 15

НОРМЫ ВЕКТОРОВ И МАТРИЦ

1. Пусть ξ—*вектор-строка*, или *точка*[1]):

$$\xi = (x_1, \ldots, x_n). \tag{15.1}$$

Мы назовем «p-нормой» вектора ξ число

$$|\xi|_p \equiv (|x_1|^p + \ldots + |x_n|^p)^{1/p} \quad (p \geqslant 1). \tag{15.2}$$

При $p \to \infty$ в скобках будет доминировать наибольший член. Пусть $|x_1| \geqslant |x_2| \geqslant \ldots \geqslant |x_n|$; тогда

$$(n|x_1|^p)^{1/p} \geqslant |\xi|_p \geqslant (|x_1|^p)^{1/p} = |x_1|.$$

С другой стороны, $\lim\limits_{p\to\infty} (n|x_1|^p)^{1/p} = |x_1|$. Поэтому удобно определить $|\xi|_\infty$ как $\max\limits_\nu |x_\nu|$:

$$|\xi|_\infty \equiv \max_\nu |x_\nu|. \tag{15.3}$$

Значения $p = 1$ и $p = \infty$ особенно важны в численном анализе. Оказывается, что в большинстве случаев в теории *сходимости* удобнее использовать $p=\infty$, в то время как $p = 1$, по-видимому, дает наилучшую норму для изучения *расходимости*.

Очевидно, что $|c\xi|_p = |c|\,|\xi|_p$.

Докажем, что

$$|\xi + \eta|_p \leqslant |\xi|_p + |\eta|_p \quad (p = 1, \infty)\,^{2)}. \tag{15.4}$$

Действительно, пусть $\eta = (y_1, \ldots, y_n)$; тогда

---

[1]) В дальнейшем мы будем употреблять оба названия.

[2]) Соотношение (15.4) справедливо при всех $p \geqslant 1$, но мы используем его только в случаях $p = 1$ и $p = \infty$.

$$|\zeta+\eta|_1=|x_1+y_1|+\ldots+|x_n+y_n|\leqslant$$
$$\leqslant|x_1|+\ldots+|x_n|+|y_1|+\ldots+|y_n|=|\xi|_1+|\eta|_1;$$
$$|\xi+\eta|_\infty=\max_\nu|x_\nu+y_\nu|\leqslant\max_\nu|x_\nu|+\max_\nu|y_\nu|=$$
$$=|\xi|_\infty+|\eta|_\infty.$$

2. Пусть $A=(a_{ij})$ — квадратная матрица порядка n. Обозначим через $|A|$ определитель матрицы A, через ξ' — *транспозицию* вектора ξ, т. е. вектор-столбец с теми же компонентами.

Мы пишем

$$A\xi'=\eta', \tag{15.5}$$

если компоненты вектора $\eta=(y_1,\ldots,y_n)$ заданы формулами

$$y_i=\sum_{j=1}^{n}a_{ij}x_j\quad(i=1,\ldots,n). \tag{15.6}$$

Из (15.6) следует, что

$$|y_i|\leqslant\sum_{j=1}^{n}|a_{ij}||x_j|\leqslant\sum_{j=1}^{n}|a_{ij}|\cdot|\xi|_\infty\quad(i=1,\ldots,n). \tag{15.7}$$

Введем теперь *норму матрицы* A формулой

$$|A|_\infty\equiv\max_i\sum_{j=1}^{n}|a_{ij}|. \tag{15.8}$$

Тогда из (15.7) получим

$$|y_i|\leqslant|A|_\infty|\xi|_\infty\quad(i=1,\ldots,n),$$

и, следовательно,

$$|\eta|_\infty\leqslant|A|_\infty|\xi|_\infty. \tag{15.9}$$

3. Покажем теперь, что в неравенстве (15.9) нельзя заменить $|A|_\infty$ на меньшую постоянную, т. е. что для любой матрицы A можно построить такой вектор ξ, который обращает (15,9) в равенство. Положим без потери общности, что $A\neq0$. Пусть

$$|A|_\infty=|a_{m1}|+|a_{m2}|+\ldots+|a_{mn}|.$$

Построим вектор ξ с компонентами

$$x_\nu = \begin{cases} \dfrac{|a_{m\nu}|}{a_{m\nu}} & \text{при } a_{m\nu} \neq 0 \\ 0 & \text{при } a_{m\nu} = 0 \end{cases} \quad (\nu = 1, \ldots, n). \tag{15.10}$$

Тогда по формуле (15.6)

$$y_m = a_{m1} x_1 + \ldots + a_{mn} x_n = \sum_{\nu=1}^{n} |a_{m\nu}| = |A|_\infty,$$

и, следовательно,

$$|\eta|_\infty \geqslant |A|_\infty. \tag{15.11}$$

При этом не все x_ν равны нулю, ибо в противном случае $A = 0$. Поэтому $|\xi|_\infty = 1$ и из (15.9) и (15.11) следует, что $|\eta|_\infty = |A|_\infty |\xi|_\infty$.

4.. Чтобы вывести аналогичные соотношения при $p = 1$, сложим полученные ранее неравенства

$$|y_i| \leqslant \sum_{j=1}^{n} |a_{ij}| |x_j| \quad (i = 1, \ldots, n). \tag{15.12}$$

Мы получим

$$|\eta|_1 \leqslant \sum_{i=1}^{n} \sum_{j=1}^{n} |a_{ij}| |x_j| = \sum_{j=1}^{n} \left(\sum_{i=1}^{n} |a_{ij}| \right) |x_j|. \tag{15.13}$$

Введем теперь *другую норму матрицы* A:

$$|A|_1 = \max_j \sum_{i=1}^{n} |a_{ij}|. \tag{15.14}$$

При этом

$$|\eta|_1 \leqslant |A|_1 \sum_{j=1}^{n} |x_j|,$$

$$|\eta|_1 \leqslant |A|_1 |\xi|_1. \tag{15.15}$$

Как и в предыдущем случае, покажем, что в этом неравенстве нельзя заменить $|A|_1$ на меньшую постоянную. Действительно, пусть $A \neq 0$ и

$$|A|_1 = \sum_{i=1}^{n} |a_{im}|.$$

В качестве ξ возьмем вектор с компонентами

$$x_m = 1; \quad x_\nu = 0 \quad (\nu \neq m).$$

Тогда

$$y_i = \sum_{j=1}^{n} a_{ij} x_j = a_{im} \quad (i = 1, \ldots, n),$$

$$|y_i| = |a_{im}|,$$

$$|\eta|_1 = \sum_{i=1}^{n} |y_i| = \sum_{i=1}^{n} |a_{im}| = |A|_1,$$

и так как $|\xi|_1 = 1$, мы получаем

$$|\eta|_1 = |A|_1 |\xi|_1.$$

5. Установим некоторые свойства норм $|A|_p$ $(p = 1, \infty)$. Очевидно, что

$$|cA|_p = |c| |A|_p \quad (p = 1, \infty). \tag{15.16}$$

Пусть $B = (b_{ij})$ — квадратная матрица того же порядка, что и A. Оценим $|A + B|_p$, где $A + B = (a_{ij} + b_{ij})$. Положим

$$\eta_1' = A\xi'; \quad \eta_2' = B\xi';$$

$$\eta_1' + \eta_2' = (A + B)\xi'.$$

Из (15.9) и (15.15) получаем

$$|\eta_1|_p \leqslant |A|_p |\xi|_p, \quad |\eta_2|_p \leqslant |B|_p |\xi|_p \quad (p = 1, \infty);$$

$$|\eta_1 + \eta_2|_p \leqslant |A + B|_p |\xi|_p \quad (p = 1, \infty). \tag{15.17}$$

Выберем вектор $\xi \neq 0$ так, чтобы в соотношении (15.17) имело место равенство. Тогда по формуле (15.4) получаем

$$|A + B|_p |\xi|_p = |\eta_1 + \eta_2|_p \leqslant |\eta_1|_p + |\eta_2|_p \leqslant$$
$$\leqslant (|A|_p + |B|_p) |\xi|_p,$$

и так как $|\xi|_p > 0$, то

$$|A + B|_p \leqslant |A|_p + |B|_p \quad (p = 1, \infty). \tag{15.18}$$

Пусть теперь $\eta' = AB\xi'$. Тогда $|\eta|_p \leqslant |A|_p |B\xi'|_p \leqslant$ $\leqslant |A|_p |B|_p |\xi|_p$, поэтому, выбирая вектор $\xi \neq 0$ так, чтобы выполнялось равенство $|\eta|_p = |AB|_p |\xi|_p$, получаем

$$|AB|_p \leqslant |A|_p |B|_p \quad (p = 1, \infty). \tag{15.19}$$

6. В дальнейшем E означает единичную матрицу, т. е. такую матрицу, у которой по главной диагонали стоят *единицы*, а на всех остальных местах — *нули*. Ясно, что $EA = AE$.

Корни λ_ν $(\nu = 1, \ldots, n)$ уравнения

$$|\lambda E - A| = 0 \tag{15.20}$$

называются *характеристическими числами* или *собственными значениями* матрицы A; уравнение (15.20) называется *характеристическим уравнением* матрицы A.

Если λ_ν — корень уравнения (15.20), то существует вектор $\xi_\nu \neq 0$, являющийся решением системы

$$(\lambda_\nu E - A)\,\xi_\nu' = 0, \quad \text{т. е. } A\xi_\nu' = \lambda_\nu \xi_\nu'. \tag{15.21}$$

ξ_ν называется *характеристическим* или *собственным вектором*, соответствующим собственному значению λ_ν. При замене A на cA каждое собственное значение умножается на c, а соответствующий собственный вектор не меняется.

Введем обозначение

$$\lambda_A = \max_\nu |\lambda_\nu| \quad (\nu = 1, \ldots, n). \tag{15.22}$$

Теорема 15.1. У любой квадратной матрицы A:

$$\lambda_A \leqslant |A|_p \quad (p = 1, \infty).$$

Доказательство. Пусть $\lambda_A = |\lambda_m|$ и ξ_m — собственный вектор, соответствующий λ_m. Тогда

$$A\xi_m' = \lambda_m \xi_m'; \quad |\lambda_m|\,|\xi_m|_p \leqslant |A|_p\,|\xi_m|_p,$$

откуда в силу $|\xi_m|_p \neq 0$ следует

$$\lambda_A = |\lambda_m| \leqslant |A|_p, \text{ ч. т. д.}$$

7. Пусть A и B — квадратные матрицы одного и того же порядка n. Говорят, что матрица B *подобна* матрице A, если существует такая *невырожденная*[1]) матрица S, что $B = SAS^{-1}$.

---

[1]) Квадратная матрица S называется невырожденной, если $S| \neq 0$. В этом случае существует обратная матрица S^{-1}, такая, что $S^{-1}S = SS^{-1} = E$; $|S^{-1}| = \frac{1}{|S|}$. — *Прим. перев.*

Теорема 15.2. Подобные матрицы имеют одни и те же собственные значения.

Действительно, у подобных матриц совпадают характеристические многочлены:

$$(\lambda E - SAS^{-1}) = S(\lambda E - A)S^{-1},$$

$$|\lambda E - SAS^{-1}| = |S||\lambda E - A||S^{-1}| = |\lambda E - A|.$$

Замечание. Если матрицы A и B подобны, то $\lambda_A = \lambda_B$. Но $|A|_p$ и $|B|_p$ могут быть различны.

8. Сформулируем теперь хорошо известную теорему Жордана (каноническая форма Жордана). *Для каждой квадратной матрицы A существует такая невырожденная матрица S, что*

$$SAS^{-1} = D + J, \tag{15.23}$$

где D — диагональная матрица, элементами которой служат собственные значения матрицы A, J — матрица с нулями и единицами на первой наддиагонали (т. е. на диагонали, параллельной главной и лежащей непосредственно над ней) *и с нулями на остальных местах.* Точнее, (15.23) может быть записано в виде

$$SAS^{-1} = \begin{pmatrix} U_1 & & & 0 \\ & U_2 & & \\ & & \ddots & \\ 0 & & & U_m \end{pmatrix}, \tag{15.24}$$

$$U_i = \begin{pmatrix} \lambda_i & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda_i & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \lambda_i & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda_i \end{pmatrix}. \tag{15.25}$$

Если λ_i — простой корень, то матрица U_i *первого* порядка. Если все собственные значения различны, то все U_i — матрицы первого порядка, J обращается в нуль и матрица A приводится преобразованием (15.23) к «диагональной форме». Это может случиться также и в случае кратных собственных значений.

9. Очевидно, что

$$|J|_p \leqslant 1 \quad (p = 1, \infty). \tag{15.26}$$

Применяя теорему Жордана к $\frac{1}{\varepsilon} A$, получаем при любом $\varepsilon \neq 0$:

$$S \frac{1}{\varepsilon} AS^{-1} = \frac{1}{\varepsilon} D + J,$$

$$SAS^{-1} = D + \varepsilon J. \tag{15.27}$$

Итак, для канонической формы Жордана несущественно, что отличные от нуля элементы матрицы J имеют величину 1; вместо 1 можно поставить любое число $\varepsilon \neq 0$. При этом по (15.26)

$$|\varepsilon J|_p = |\varepsilon| \, |J|_p \leqslant \varepsilon. \tag{15.28}$$

Теорема 15.3. Для любого $\varepsilon > 0$ найдется такая невырожденная матрица S, что

$$\lambda_A \leqslant |SAS^{-1}|_p \leqslant \lambda_A + \varepsilon \quad (p = 1, \infty). \tag{15.29}$$

Действительно, из теоремы 15.1 и неравенства (15.28) имеем

$$\lambda_A \leqslant |SAS^{-1}|_p \leqslant |D|_p + |\varepsilon J|_p = \lambda_A + \varepsilon.$$

Так как собственные значения матрицы A являются корнями алгебраического уравнения (15.20), то к ним применимы результаты приложений I и II. Однако прямое вычисление коэффициентов характеристического уравнения сопряжено с большими трудностями. Поэтому важно иметь оценки вариации корней уравнения (15.20) в зависимости от вариации элементов матрицы A. Такие оценки даны в приложении XI.

ГЛАВА 16

ДВЕ ТЕОРЕМЫ О СХОДИМОСТИ ПРОИЗВЕДЕНИЙ МАТРИЦ

1. Следующая теорема дает оценку нормы произведений вида

$$\Pi_m = \prod_{\mu=1}^{m} (A + U_\mu), \tag{16.1}$$

где A и U_μ $(\mu = 1, \ldots, m)$ — квадратные матрицы одного порядка.

Теорема 16.1. Пусть заданы матрица A и число $\varepsilon > 0$. Тогда существуют положительные постоянные $\eta_1 > 0$ и $\sigma > 0$, зависящие только от A и ε, такие, что если

$$|U_\mu|_\infty \leqslant \eta_1 \quad (\mu = 1, 2, \ldots), \tag{16.2}$$

то, независимо от порядка сомножителей в (16.1),

$$|\Pi_m|_\infty \leqslant \sigma (\lambda_A + \varepsilon)^m \quad (m = 1, 2, \ldots). \tag{16.3}$$

В частности, если $\lambda_A < 1$, то существует $\eta_1 > 0$, при котором

$$\Pi_m \to 0 \quad (m \to \infty). \tag{16.4}$$

2. Доказательство. По теореме 15.3 найдется матрица S, для которой имеет место соотношение

$$|SAS^{-1}|_\infty \equiv s < \lambda_A + \frac{\varepsilon}{2}. \tag{16.5}$$

Определим σ и η_1 формулами

$$\sigma = |S|_\infty |S^{-1}|_\infty, \quad \eta_1 = \frac{\varepsilon}{2\sigma} \tag{16.6}$$

и положим

$$B = SAS^{-1}; \quad V_\mu = SU_\mu S^{-1} \quad (\mu = 1, 2, \ldots). \tag{16.7}$$

Тогда

$$S\Pi_m S^{-1} = [S(A+U_1)S^{-1}]\ldots[S(A+U_m)S^{-1}] =$$
$$= [SAS^{-1}+SU_1S^{-1}]\ldots[SAS^{-1}+SU_mS^{-1}],$$
$$S\Pi_m S^{-1} = \prod_{\mu=1}^{m}(B+V_\mu). \qquad (16.8)$$

Из (16.2) и (16.5) — (16.8) следует, что

$$|V_\mu|_\infty \leqslant \sigma\eta_1 = \frac{\varepsilon}{2}; \quad |B+V_\mu|_\infty \leqslant s+\sigma\eta_1 < \lambda_A+\varepsilon,$$

$$|S\Pi_m S^{-1}|_\infty \leqslant (\lambda_A+\varepsilon)^m. \qquad (16.9)$$

Так как при этом $|\Pi_m|_\infty = |S^{-1}(S\Pi_m S^{-1})S|_\infty$, то из (16.9) и (16.6) получаем

$$|\Pi_m|_\infty \leqslant \sigma(\lambda_A+\varepsilon)^m, \text{ ч. т. д.}$$

Если $\lambda_A < 1$, то, выбирая такое $\varepsilon > 0$, чтобы было $\lambda_A + \varepsilon < 1$, и соответствующее $\eta_1 > 0$, получаем (16.4) как следствие (16.3).

3. Теорема 16.2. Пусть A — квадратная матрица с $\lambda_A < 1$, и пусть $\varepsilon > 0$ таково, что $\lambda_A + \varepsilon < 1$. Построим последовательность матриц A_μ по формулам

$$A_1 = A; \quad A_{\mu+1} = A_\mu A + W_\mu. \qquad (16.10)$$

Существуют положительные постоянные $\eta_2 > 0$ и $\sigma > 0$, зависящие только от A и ε и такие, что если

$$|W_\mu|_\infty \leqslant \eta_2 |A_\mu|_\infty \quad (\mu = 1,2,\ldots), \qquad (16.11)$$

то

$$|A_\mu|_\infty \leqslant \sigma(\lambda_A+\varepsilon)^\mu \quad (\mu = 1,2,\ldots), \quad A_\mu \to 0 \quad (\mu\to\infty). \qquad (16.12)$$

4. Доказательство. Пусть S, s, B и σ имеют тот же смысл, что и при доказательстве теоремы 16.1. Определим η_2, T_μ и B_μ следующим образом:

$$\eta_2 = \frac{\varepsilon}{2\sigma^2}, \qquad (16.13)$$

$$T_\mu = SW_\mu S^{-1}; \quad B_\mu = SA_\mu S^{-1} \quad (\mu = 1,2,\ldots), \quad B_1 = B. \qquad (16.14)$$

По формуле (16.10)

$$SA_{\mu+1}S^{-1} = SA_\mu S^{-1} SAS^{-1} + SW_\mu S^{-1}$$

и, следовательно,

$$B_{\mu+1} = B_\mu B + T_\mu \quad (\mu = 1, 2, \ldots). \tag{16.15}$$

С другой стороны, из (16.14), (16.6) и (16.11) следует, что

$$\begin{gathered} |T_\mu|_\infty \leqslant |S|_\infty |S^{-1}|_\infty |W_\mu|_\infty, \\ |T_\mu|_\infty \leqslant \sigma\eta_2 |A_\mu|_\infty. \end{gathered} \tag{16.16}$$

5. Но из (16.14) мы имеем $A_\mu = S^{-1} B_\mu S$, и, значит,

$$|A_\mu|_\infty \leqslant \sigma \ |B_\mu|_\infty \quad (\mu = 1, 2, \ldots). \tag{16.17}$$

Так как $|B|_\infty = s$, то из (16.15), (16.16) и (16,17) вытекает

$$|B_{\mu+1}|_\infty \leqslant s|B_\mu|_\infty + \sigma^2 \eta_2 |B_\mu|_\infty,$$

$$|B_{\mu+1}|_\infty \leqslant |B_\mu|_\infty \left(s + \frac{\varepsilon}{2}\right) \quad (\mu = 1, 2, \ldots),$$

$$|B_{\mu+1}|_\infty \leqslant (\lambda_A + \varepsilon) |B_\mu|_\infty \leqslant (\lambda_A + \varepsilon)^\mu |B|_\infty \to 0 \quad (\mu \to \infty).$$

Отсюда на основании формул (16.5), (16.6), (16.7) и (16.17) получаем (16.12).

6. Теоремы 16.1 и 16.2 дают обобщение известного факта, что μ-я степень матрицы A стремится к нулю, если $\lambda_A < 1$. В частности, из второй теоремы вытекает *устойчивость сходимости* степени A^μ к 0 относительно округления на каждом этапе вычисления.

ГЛАВА 17

ОДНА ТЕОРЕМА О РАСХОДИМОСТИ ПРОИЗВЕДЕНИЙ МАТРИЦ

1. Докажем теперь теорему, соответствующую теореме 16.1 при $\lambda_A > 1$. При доказательстве мы используем норму с $p = 1$.

Теорема 17.1. Пусть A — квадратная матрица порядка n с $\lambda_A > 1$, и пусть $\varepsilon > 0$ таково, что $\lambda_A - \varepsilon > 1$. Построим для последовательности U_μ $(\mu = 1, 2, \ldots)$ квадратных матриц того же порядка n произведения

$$\Pi_m \equiv \prod_{\mu=1}^{m} (A + U_\mu) \quad (m = 1, 2, \ldots). \tag{17.1}$$

Существует такое $\delta = \delta(A, \varepsilon) > 0$, что если $|U_\mu|_1 \leqslant \delta$ $(\mu = 1, 2, \ldots)$, то произведение (17.1) расходится при $m \to \infty$. Точнее, в этом случае существует телесный угол L с вершиной в начале координат, такой, что для любого вектора $\zeta \neq 0$, лежащего в L,

$$(\lambda_A - \varepsilon)^{-m} |\Pi_m \zeta'|_1 \to \infty \quad (m \to \infty), \tag{17.2}$$

независимо от порядка сомножителей в Π_m.

2. Доказательство. Пусть собственные значения $\lambda_1, \lambda_2, \ldots, \lambda_n$ матрицы A перенумерованы так, что

$$\lambda_A = |\lambda_1| \geqslant |\lambda_2| \geqslant \ldots \geqslant |\lambda_k| = t > 1 \geqslant |\lambda_{k+1}| \geqslant \ldots \geqslant |\lambda_n|. \tag{17.3}$$

Положим

$$\gamma = \frac{t-1}{2n+1} \tag{17.4}$$

и выберем некоторое τ, $0 < \tau < \gamma$. Приведем матрицу A к канонической форме Жордана вида (15.27):

$$SAS^{-1} = B = D + \tau J. \tag{17.5}$$

Положим

$$\sigma = |S|_1 |S^{-1}|_1; \quad \delta = \frac{\gamma - \tau}{\sigma}; \tag{17.6}$$

$$V_\mu = SU_\mu S^{-1}; \quad W_\mu = V_\mu + \tau J. \tag{17.7}$$

При этом в силу (17.5)

$$B + V_\mu = D + W_\mu. \tag{17.8}$$

3. Ясно, что

$$\begin{gathered} S(A + U_\mu) S^{-1} = SAS^{-1} + SU_\mu S^{-1} = B + V_\mu, \\ S\left[\prod_{\mu=1}^{m} (A + U_\mu)\right] S^{-1} = \prod_{\mu=1}^{m} (D + W_\mu). \end{gathered} \tag{17.9}$$

Так как по условию $|U_\mu|_1 \leqslant \delta$, то из (17.6) следует, что

$$|V_\mu|_1 \leqslant |S|_1 |U_\mu|_1 |S^{-1}|_1 \leqslant \sigma\delta = \gamma - \tau.$$

Поэтому

$$|W_\mu|_1 \leqslant \gamma. \tag{17.10}$$

Таким образом, достаточно доказать нашу теорему для диагональной матрицы D.

Для любого вектора ζ имеем

$$S\left[\prod_{\mu=1}^{m} (A + U_\mu)\right] S^{-1} (S\zeta') = \left[\prod_{\mu=1}^{m} (D + W_\mu)\right] S\zeta'. \tag{17.11}$$

Мы докажем, что если компоненты вектора $\xi = \xi_0 \neq 0$ удовлетворяют неравенству [1])

$$|x_1| + |x_2| + \ldots + |x_k| \geqslant |x_{k+1}| + \ldots + |x_n|, \tag{17.12}$$

то

$$\left| \prod_{\mu=1}^{m} (D + W_\mu) \xi_0' \right|_1 \to \infty \quad (m \to \infty). \tag{17.13}$$

---

[1]) В случае $k=n$ в качестве ξ_0 можно брать любой вектор $\xi \neq 0$; при этом дальнейшее доказательство упрощается, так как $\beta_m = 0$; $\alpha_m = |\xi_m|_1 > 0$ и мы сразу получаем (17.25) в виде $\alpha_{m+1} \geqslant [1 + \gamma(n+1)]\alpha_m$, откуда следует (17.27). — *Прим. перев.*

4. Положим

$$\xi'_m = \prod_{\mu=1}^{m} (D + W_\mu)\,\xi'_0 \quad (m = 1, 2, \ldots), \tag{17.14}$$

$$\xi_m = \left(x_1^{(m)}, \ldots, x_n^{(m)}\right) \quad (m = 0, 1, \ldots). \tag{17.15}$$

При этом

$$\xi'_{m+1} = (D + W_{m+1})\xi'_m \quad (m = 0, 1, \ldots). \tag{17.16}$$

Положим, далее,

$$\alpha_m = \left|x_1^{(m)}\right| + \ldots + \left|x_k^{(m)}\right|, \tag{17.17}$$

$$\beta_m = \left|x_{k+1}^{(m)}\right| + \ldots + \left|x_n^{(m)}\right|, \tag{17.18}$$

$$|\xi_m|_1 = \alpha_m + \beta_m \quad (m = 0, 1, \ldots). \tag{17.19}$$

5. Из формулы (17.12):

$$\alpha_0 \geqslant \beta_0. \tag{17.20}$$

Докажем по индукции, что для всех $m \geqslant 0$:

$$\alpha_m \geqslant \beta_m \quad (m = 0, 1, \ldots). \tag{17.21}$$

Пусть неравенство (17.21) справедливо при некотором $m \geqslant 0$. Полагая $W_{m+1} = (w_{lj})$, запишем (17.16) в координатной форме

$$x_l^{(m+1)} = \lambda_l x_l^{(m)} + \sum_{j=1}^{n} w_{lj} x_j^{(m)}. \tag{17.22}$$

Но из (17.10), в частности, следует, что

$$|w_{lj}| \leqslant \gamma \quad (l, j = 1, \ldots, n)$$

и потому

$$\left|\sum_{j=1}^{n} w_{lj} x_j^{(m)}\right| \leqslant \gamma(\alpha_m + \beta_m). \tag{17.23}$$

Следовательно,

$$\left|x_l^{(m+1)}\right| \geqslant |\lambda_l| \left|x_l^{(m)}\right| - \gamma(\alpha_m + \beta_m) \quad (l = 1, \ldots, n). \tag{17.24}$$

6. Из (17.3) и (17.21) при $l = 1, \ldots, k$ имеем

$$\left|x_l^{(m+1)}\right| \geqslant t \left|x_l^{(m)}\right| - 2\gamma\alpha_m.$$

Суммируя по l от 1 до k, получаем

$$\alpha_{m+1} \geqslant t\alpha_m - 2k\gamma\alpha_m = (t - 2k\gamma)\,\alpha_m,$$

или, так как $t = 1 + (2n+1)\,\gamma$,

$$\alpha_{m+1} \geqslant [1 + \gamma + 2\,(n-k)\,\gamma]\,\alpha_m. \tag{17.25}$$

С другой стороны, из (17.3), (17.23), и (17.21) при $l = k+1, \ldots, n$ получаем

$$\begin{gathered} |x_l^{(m+1)}| \leqslant |x_l^{(m)}| + 2\gamma\alpha_m, \\ \beta_{m+1} \leqslant [1 + 2\,(n-k)\,\gamma]\,\alpha_m. \end{gathered} \tag{17.26}$$

Из (17.25) и (17.26) следует

$$\alpha_{m+1} > \beta_{m+1}.$$

Таким образом, неравенство (17.21) *справедливо для всех* $m = 0, 1, \ldots$.

7. Поэтому (17.25) также справедливо для всех $m = 0, 1, \ldots$. Но при этом

$$\alpha_{m+1} \geqslant (1+\gamma)\,\alpha_m \quad (m = 0, 1, \ldots),$$

и, следовательно,

$$\begin{gathered} \alpha_m \to \infty \quad (m \to \infty), \\ |\xi_m|_1 \to \infty \quad (m \to \infty), \end{gathered} \tag{17.27}$$

что и доказывает (17.13).

Вернемся теперь к матрице A. Если вектор $\zeta' = S^{-1}\xi'$ лежит в телесном угле, получающемся из (17.12) преобразованием S^{-1}, то для вектора ξ справедливо (17.13), следовательно,

$$|\Pi_m \zeta'|_1 \to \infty.$$

Чтобы закончить доказательство теоремы 17.1, рассмотрим матрицы

$$C = \frac{1}{\lambda_A - \varepsilon} A; \qquad X_\mu = \frac{1}{\lambda_A - \varepsilon} U_\mu.$$

Здесь

$$\lambda_c = \frac{\lambda_A}{\lambda_A - \varepsilon} > 1.$$

Поэтому существуют положительное число δ_ε и телесный угол L с вершиной в начале координат, такие, что если $|X_\mu|_1 \leqslant \delta_\varepsilon$, то для любого вектора ζ из L:

$$(\lambda_A - \varepsilon)^{-m} |\Pi_m \zeta'|_1 = \left| \prod_{\mu=1}^{m} (C + X_\mu) \zeta' \right|_1 \to \infty.$$

Последнее соотношение справедливо, если

$$|U_\mu|_1 \leqslant (\lambda_A - \varepsilon)\, \delta_\varepsilon = \delta(A, \varepsilon) \quad (\mu = 1, 2, \ldots).$$

Теорема доказана.

8. Следствие. При условиях теоремы (17.1) мы имеем

$$\frac{|\Pi_m|_1}{|\lambda_A - \varepsilon|^m} \to \infty \quad (m \to \infty). \tag{17.28}$$

ГЛАВА 18

ХАРАКТЕРИСТИКА ТОЧЕК ПРИТЯЖЕНИЯ И ОТТАЛКИВАНИЯ ПРИ ИТЕРАЦИЯХ С НЕСКОЛЬКИМИ ПЕРЕМЕННЫМИ

1. Пусть $\xi = (x_1, \ldots, x_n)$ — точка в n-мерном вещественном пространстве; положим

$$y_i = f_i(x_1, \ldots, x_n) \equiv f_i(\xi) \quad (i = 1, \ldots, n). \tag{18.1}$$

Применяя векторные обозначения, мы будем записывать (18.1) в виде

$$\eta = \Phi(\xi), \tag{18.2}$$

где

$$\eta = (y_1, \ldots, y_n) = [f_1(\xi), \ldots, f_n(\xi)]. \tag{18.3}$$

Точка ζ называется *неподвижной точкой* или *центром* преобразования (18.2), если

$$\zeta = \Phi(\zeta). \tag{18.4}$$

Пусть теперь ξ_0 — «начальное приближение» к центру ζ; мы построим последовательность приближений $\{\xi_k\}$ к центру ζ с помощью итерации

$$\xi_1 = \Phi(\xi_0), \ldots, \quad \xi_{k+1} = \Phi(\xi_k), \ldots . \tag{18.5}$$

Если существует такая окрестность точки ζ, что для каждого ξ_0 из этой окрестности последовательность (18.5) сходится к ζ, то ζ называется *точкой притяжения*, в противном случае — *точкой отталкивания*.

2. Обозначим матрицу Якоби преобразования (18.2) через

$$\Delta(\xi) = \left(\frac{\partial f_i(\xi)}{\partial x_j}\right)_j, \tag{18.6}$$

где i — индекс строки, j — индекс столбца; обозначим, далее,

$$\Delta(\zeta) = A. \tag{18.7}$$

Теорема 18.1. Пусть функции $f_i(\xi)$, $(i = 1, \ldots, n)$, дифференцируемы в некоторой выпуклой окрестности V^* центра ζ и их первые производные непрерывны в самой точке ζ. Если для матрицы (18.7)

$$\lambda_A < 1, \tag{18.8}$$

то ζ есть точка притяжения, точнее существуют такие две ее окрестности V, V_0, что для каждой точки $\xi_0 \in V$ последовательность $\{\xi_k\}$, определенная итерацией (18.5), содержится в V_0 и

$$\xi_k \to \zeta \quad (k \to \infty). \tag{18.9}$$

3. Доказательство. Без ограничения общности можно предположить, что ζ совпадает с началом координат. Так как при этом $f_i(0) = 0$ $(i = 1, \ldots, n)$, то, применяя теорему о среднем значении, получим для $\xi \in V^*$:

$$y_i = f_i(\xi) - f_i(0) = \sum_{j=1}^{n} x_j \frac{\partial f_i(\vartheta_i \xi)}{\partial x_j}, \quad |\vartheta_i| \leqslant 1 \quad (i=1, 2, \ldots n). \tag{18.10}$$

Эти равенства можно записать в векторной форме

$$\eta' = A(\xi)\,\xi', \tag{18.11}$$

введя матрицу

$$A(\xi) = \left(\frac{\partial f_i(\vartheta_i \xi)}{\partial x_j}\right). \tag{18.12}$$

Введем еще матрицу $U(\xi) = (u_{ij})$ по формуле:

$$A(\xi) = A + U(\xi), \tag{18.13}$$

$$u_{ij} = \frac{\partial f_i(\vartheta_i \xi)}{\partial x_j} - \frac{\partial f_i(0)}{\partial x_j}. \tag{18.14}$$

4. Так как первые производные от $f_i(\xi)$ непрерывны в ζ, мы можем сделать (u_{ij}) сколь угодно малыми, выбирая вектор ξ с достаточно малой нормой $|\xi|_\infty$. Возьмем такое $\varepsilon > 0$, для которого $\lambda_A + \varepsilon < 1$. Найдем η_1 и σ, соответствующие A и ε по теореме 16.1, и выберем такое $\tau > 0$, чтобы выпуклая окрестность V_0 начала координат, определяемая неравенством

$$|\xi|_\infty \leqslant \tau,$$

лежала в V^* и чтобы всюду в V_0 выполнялись неравенства

$$|u_{ij}| < \frac{\eta_1}{n+1} \quad (i,j = 1, \ldots, n). \tag{18.15}$$

Наконец, определим окрестность V формулой

$$V = \frac{1}{1+\sigma} V_0. \tag{18.16}$$

Докажем, что если $\xi_0 \in V$, то все $\xi_k \in V_0$. Предположим, что при некотором $k \geqslant 0$ выполнены условия

$$\xi_0 \in V \subset V_0; \quad \xi_1, \ldots, \xi_k \in V_0.$$

Тогда по формулам (18.5), (18.11) и (18.13) имеем

$$\xi'_{k+1} = [A + U(\xi_k)]\, \xi'_k,$$

откуда

$$\xi'_{k+1} = \prod_{\nu=0}^{k} [A + U(\xi_\nu)]\, \xi'_0. \tag{18.17}$$

5. Из (18.15) следует, что

$$|U(\xi_\nu)|_\infty < \eta_1 \quad (\nu = 0, 1, \ldots, k),$$

поэтому мы можем применить теорему 16.1. По формуле (16.3) имеем

$$\left| \prod_{\nu=0}^{k} [A + U(\xi_\nu)] \right|_\infty \leqslant \sigma (\lambda_A + \varepsilon)^{k+1} < \sigma,$$

следовательно, по (18.17)

$$|\xi_{k+1}|_\infty \leqslant \sigma |\xi_0|_\infty. \tag{18.18}$$

Учитывая формулу (18.16), мы находим отсюда, что

$$\xi_{k+1} \in \frac{\sigma}{1+\sigma} V_0 \subset V_0,$$

так что наше предположение верно также и для ξ_{k+1}, а значит, и для всех ξ_k. Но тогда при всех k

$$\xi'_{k+1} = \prod_{\nu=0}^{k} [A + U(\xi_\nu)]\, \xi'_0 \quad (k = 0, 1, \ldots)$$

и, так как $\lambda_A < 1$, то по (16.4)

$$\xi_{k+1} \to 0 \quad (k \to \infty), \text{ ч. т. д.} \tag{18.19}$$

6. Теорема 18.2. Пусть функции $f_i(\xi)$ $(i = 1, \ldots, n)$ дифференцируемы в некоторой окрестности V^* центра ζ, и пусть их первые производные непрерывны в самой точке ζ. Если для матрицы (18.7)

$$\lambda_A > 1, \tag{18.20}$$

то ζ есть точка отталкивания; точнее, существуют такая окрестность V центра ζ и такой телесный угол L с вершиной в ζ, что при любой начальной точке $\xi_0 \in LV$ (LV — пересечение областей L и V) у последовательности (18.5) найдется либо элемент, совпадающий с ζ, либо элемент, лежащий вне V.

7. Доказательство. Без ограничения общности можно предположить, что ζ совпадает с началом координат. Возьмем такое $\varepsilon > 0$, для которого $\lambda_A - \varepsilon > 1$. Найдем $\delta > 0$ и телесный угол L с вершиной в начале координат, для которых справедливы утверждения теоремы 17.1, и определим окрестность V начала координат так, чтобы в любой точке $\xi \in V$ было

$$\left| \frac{\partial f_i(\xi)}{\partial x_j} - \frac{\partial f_i(0)}{\partial x_j} \right| < \frac{\delta}{n+1}. \tag{18.21}$$

Предположим теперь, что вопреки утверждению теоремы найдется последовательность (18.5), начальная точка которой ξ_0 принадлежит LV, а все остальные точки ξ_ν лежат внутри V, не совпадая с началом координат. Определив $A(\xi_\nu)$ и $U(\xi_\nu)$ по формулам (18.12) и (18.13), получаем в силу (18.21)

$$|U(\xi_\nu)|_1 < \delta \quad (\nu = 0, 1, \ldots),$$

и наше предположение приходит в противоречие с теоремой 17.1. Теорема 18.2 доказана.

8. Поскольку точка ζ нам неизвестна, мы должны прежде всего проверить, будет ли λ_A меньше единицы. На практике обычно приходится доказывать, что всюду внутри рассматриваемой области

$$\lambda_{\Delta(\xi)} < 1. \tag{18.22}$$

Доказательство этого неравенства может оказаться весьма затруднительным, так как мы имеем дело с корнями

многочлена n-й степени, причем такие многочлены надо исследовать в каждой точке области. Однако в данном случае можно применять различные оценки собственных чисел матриц. Простейшие оценки такого рода дает теорема 15.1.

Если, например, при некотором положительном $q < 1$ имеют место неравенства

$$\sum_{j=1}^{n} \left| \frac{\partial f_i(\xi)}{\partial x_j} \right| \leqslant q \quad (i = 1, \ldots, n) \tag{18.23}$$

во всех точках ξ из некоторой окрестности центра ζ, то по теореме 15.1 во всех точках этой окрестности

$$\lambda_{\Delta(\xi)} \leqslant |\Delta(\xi)|_\infty \leqslant q < 1. \tag{18.24}$$

ПРИЛОЖЕНИЯ

ПРИЛОЖЕНИЕ I

НЕПРЕРЫВНОСТЬ КОРНЕЙ АЛГЕБРАИЧЕСКИХ УРАВНЕНИЙ

1. Разрешая уравнения, содержащие числовые параметры, мы обычно вынуждены округлять значения этих параметров, т. е. заменять их приближенными значениями. О влиянии такой замены на величину корня в общем случае можно сказать очень немного. Конечно, если уравнение с неизвестным z и параметром t имеет вид $f(z, t) = 0$, то влияние вариации параметра t можно учесть с помощью соотношения $dz/dt = -f'_t/f'_z$. Однако это соотношение можно применять лишь после того, как величины z и f'_z вычислены или хотя бы оценены с достаточной точностью.

Результаты общего характера получены лишь для алгебраических уравнений. Эти результаты очень слабы именно из-за своей большой общности.

2. Рассмотрим два многочлена

$$\begin{aligned} p(z) &= a_0 z^n + \ldots + a_n, \qquad a_0 = 1; \\ q(z) &= b_0 z^n + \ldots + b_n, \qquad b_0 = 1. \end{aligned} \tag{I.1}$$

Пусть $x_1, \ldots, x_n$ — корни многочлена $p(z)$; $y_1, \ldots, y_n$ — корни многочлена $q(z)$. Мы хотим получить для разностей между x_ν и y_ν оценки, зависящие только от величин $|b_\nu - a_\nu|$.

Положим

$$\gamma = \max_\nu (|x_\nu|, \ |y_\nu|), \ \Gamma = \max_{\nu > 0} (|a_\nu|^{1/\nu}, \ |b_\nu|^{1/\nu}). \tag{I.2}$$

Известно, что $\gamma \leqslant 2\Gamma$[1]).

---

[1]) Эта оценка может быть улучшена благодаря результату, полученному еще Лагранжем; но в настоящее время почти забытому (Lagrange J. L, De la résolution des équations numériques, Paris, Duprat, VI (1797/8), стр. 16; перепечатано в Oeuvres complètes, том VIII, стр. 32). Если числа $|a_\nu|^{1/\nu}$ расположены в порядке возрастания и сумма двух последних чисел этого ряда равна L_p, то все $|x_\nu| \leqslant L_p$. В нашем случае отсюда следует, что $\gamma \leqslant \max(L_p, L_q)$.

Положим

$$\varepsilon = \sqrt[n]{\sum_{\nu=1}^{n} |b_\nu - a_\nu| \gamma^{n-\nu}}. \tag{I.3}$$

Так как

$$q(z) - p(z) = \sum_{\nu=1}^{n} (b_\nu - a_\nu) z^{n-\nu},$$

то, обозначая через x_0 любой из корней $x_1, \ldots, x_n$, получаем

$$|q(x_0)| = \prod_{\nu=1}^{n} |x_0 - y_\nu| \leqslant \varepsilon^n. \tag{I.4}$$

3. Мы видим, что для выбранного нами x_0 всегда найдется один из корней (скажем y_0), для которого

$$|x_0 - y_0| \leqslant \varepsilon. \tag{I.5}$$

Так как в предыдущем рассуждении $p(z)$ и $q(z)$ можно поменять местами, то *в ε-окрестности любого корня одного из многочленов $p(z)$ или $q(z)$ всегда лежит корень другого многочлена.* Это не означает, однако, что можно так перенумеровать x_ν и y_ν, чтобы при любом ν выполнялось неравенство $|x_\nu - y_\nu| \leqslant \varepsilon$; последние неравенства, действительно, могут не иметь места, если $n > 2$ (см. пример в конце настоящего приложения, п. 10).

4. Однако справедлива следующая

Теорема. Корни x_ν, y_ν можно перенумеровать так, чтобы выполнялись неравенства

$$|x_\nu - y_\nu| < 2n\varepsilon \quad (\nu = 1, \ldots, n). \tag{I.6}$$

При доказательстве этой теоремы можно считать, что указанные выше ε-окрестности корней x_ν и y_ν *открыты*, т. е. *представляют собой внутренности кругов*. Действительно, из формулы (I.4) следует, что если для некоторого корня x_0 многочлена $p(z)$ не существует такого y_ν, что $|x_0 - y_\nu| < \varepsilon$, то для всех y_ν должны выполняться равенства $|x_0 - y_\nu| = \varepsilon$ ($\nu = 1, \ldots, n$). Но в этом специальном случае справедливость неравенства (I.6) при $n > 1$ вытекает из того, что в ε-окрестности любого x_ν лежит один из корней y_μ, и поэтому расстояние от любого x_ν до любого y_μ не превосходит $3\varepsilon < 4\varepsilon$.

5. Прежде чем доказывать теорему, мы докажем следующую лемму из теории функций комплексного переменного.

Лемма. Пусть в z-плоскости дана замкнутая область G, граница которой состоит из конечного числа гладких дуг. Пусть функции $f(z)$, $h(z)$ регулярны в G. Если ни при каком значении действительного параметра t из сегмента $a \leqslant t \leqslant b$ функция $f(z) + th(z)$ не обращается в нуль на границе области G, то число $N(t)$ нулей этой функции, лежащих внутри области G, не зависит от t при $a \leqslant t \leqslant b$.

Доказательство леммы. Пусть t_0 — любое значение из $[a, b]$. Так как функция $u(z) = f(z) + t_0 h(z)$ не обращается в нуль на границе области G, то ее модуль имеет положительную нижнюю грань p на границе области.

Выберем $|\delta|$ настолько малым, чтобы на границе области G модуль функции $v(z) = \delta h(z)$ был меньше p и, значит, $|v| < |u|$. Тогда по теореме Руше функции $u + v = f(z) + (t_0 + \delta) h(z)$ и $u = f(z) + t_0 h(z)$ имеют одинаковое число нулей внутри области G, т. е. $N(t_0 + \delta) = N(t_0)$. Это означает, что функция $N(t)$ постоянна в некоторой окрестности любого $t_0 \in [a, b]$ и, следовательно, непрерывна на $[a, b]$. Но так как функция $N(t)$ может принимать только целочисленные значения, то она постоянна на всем отрезке $[a, b]$. Лемма доказана.

6. Доказательство теоремы. Назовем два корня x_ν и x_μ многочлена $p(z)$ *соседними*, если расстояние между ними не превосходит 2ε. Два корня x_ν и x_μ многочлена $p(z)$ назовем *связанными*, если существует последовательность корней многочлена $p(z)$, начинающаяся с x_ν и кончающаяся на x_μ: $x_\nu, x_{\lambda_1}, x_{\lambda_2}, \ldots, x_{\lambda_\varkappa}, x_\mu$, в которой любые два рядом стоящих корня являются соседними.

Все n корней многочлена $p(z)$ могут быть разбиты на несколько групп: $g_1, \ldots, g_k$ таким образом, что все корни внутри одной группы связаны между собой, а два корня, принадлежащие к различным группам, не могут быть связаны.

Обозначим объединение открытых ε-окрестностей корней, принадлежащих $g_\varkappa$, через $G_\varkappa$; границу открытой об-

ласти $G_\varkappa$ обозначим через $B_\varkappa$. $B_\varkappa$ состоит из конечного числа дуг окружностей. $G_\varkappa$ и G_λ не имеют общих точек, если $\varkappa \neq \lambda$.

7. Предположим теперь, что группа g_1 состоит из ϱ корней, считаемых, конечно, вместе с кратностью. Рассмотрим функцию

$$q_t(z) = p(z) + t\,[q(z) - p(z)] \quad (0 \leqslant t \leqslant 1). \tag{I.7}$$

Заметим, что ни при одном t из сегмента $[0, 1]$ функция $g_t(z)$ не имеет корней на B_1. Действительно, такой корень должен был бы лежать на расстоянии $\geqslant \varepsilon$ от всех корней $p(z)$, в то время как из рассуждений п. 1—4 следует, что хотя бы один из корней x_ν находится от этого корня на расстоянии $\leqslant \varepsilon t^{1/n}$ при $t < 1$ и на расстоянии $< \varepsilon$ при $t = 1$.

Из нашей леммы следует теперь, что функция (I.7) имеет при $0 \leqslant t \leqslant 1$ постоянное число корней внутри G_1. При $t = 0$ это число есть ϱ, и мы видим, что $q(z)$ также имеет в области G_1 точно ϱ корней y_ν. Так как при этом расстояние от любого из этих y_ν до любого корня x_ν из g_1 не превосходит $(2\varrho - 1)\,\varepsilon < 2n\varepsilon$, то (I.6) доказано для x_ν, y_ν, содержащихся в G_1.

Очевидно, те же рассуждения можно провести для любой группы $g_\varkappa$. Теорема доказана.

8. Для применения (I.6) необходимо, разумеется, получить удобные оценки величины ε. Положим

$$\delta = \max_\nu |\, b_\nu - a_\nu \,|. \tag{I.8}$$

Так как $\gamma \leqslant 2\Gamma$, то из (I.3) следует

$$\varepsilon^n \leqslant \delta \sum_{\nu=0}^{n-1} (2\Gamma)^\nu.$$

С другой стороны, при любом $u \geqslant 0$ имеет место неравенство[1)]

[1)] В справедливости неравенства (I.9) нетрудно убедиться, рассматривая порознь случаи $u < 1$ и $u > 1$. При этом неравенство с n устанавливается непосредственно, а неравенство с $\frac{1}{|1-u|}$ доказывается заменой выражения в левой части (I.9) на $\frac{1-u^n}{1-u}$.

$$\sum_{\nu=0}^{n-1} u^\nu \leqslant \max(1,\ u^n)\cdot\min\left(n,\ \frac{1}{|1-u|}\right). \tag{I.9}$$

Это приводит к следующей оценке:

$$\varepsilon \leqslant \delta^{\frac{1}{n}} \max(1,\ 2\Gamma)\sqrt[n]{\min\left(n,\ \frac{1}{|1-2\Gamma|}\right)}. \tag{I.10}$$

Неравенство (I.10) может оказаться неудобным из-за того, что все разности $b_\nu - a_\nu$ входят в него с одинаковым *весом*. Можно получить другую оценку, предположив, что

$$|b_\nu - a_\nu| \leqslant \sigma\Gamma^\nu \qquad (\nu = 1, \ldots, n). \tag{I.11}$$

При этом по (I.3)

$$\varepsilon^n \leqslant \sigma \sum_{\nu=1}^{n} \Gamma^\nu (2\Gamma)^{n-\nu} = \sigma\Gamma^n (1 + 2 + \ldots + 2^{n-1}),$$

$$\varepsilon \leqslant 2\Gamma\sigma^{1/n}. \tag{I.12}$$

9. Очень поучительный пример дает уравнение

$$(z-1)^4 - (7\cdot 10^{-4} z)^2 = z^4 - 4z^3 + (6 - 49\cdot 10^{-8}) z^2 -$$

$$- 4z + 1 = 0, \tag{I.13}$$

имеющее четыре корня:

$$y_1 = 1{,}02681; \quad y_2 = 0{,}97389;$$

$$y_{3,4} = 0{,}99965 \pm i\,0{,}026455.$$

Возьмем многочлен (I.13) за $q(z)$ и многочлен $(z-1)^4$ за $p(z)$. Тогда все x_ν равны 1,

$$\max_\nu |y_\nu - x_\nu| = 0{,}02681 = y_1 - 1,$$

и

$$\varepsilon^4 = 49\cdot 10^{-8} y_1^2 = (y_1 - 1)^4, \quad \varepsilon = y_1 - 1.$$

Поэтому в данном случае имеет место знак равенства в (I.5).

10. Покажем, наконец, на примере, что в случае $n=3$ нельзя в оценке (I.6) заменить коэффициент $2n$ на 1. Возьмем

$$p(z) = z^3 + \frac{3}{2}\sqrt{2}\,z^2 - 1; \quad q(z) = z^3 + \frac{3}{2}\sqrt[3]{2}\,z^2;$$

здесь

$$x_1 = -\sqrt[3]{2}; \quad x_2 = -\sqrt[3]{2}; \quad x_3 = \frac{1}{\sqrt[3]{4}} = 0{,}62996;$$

$$y_1 = 0; \quad y_2 = 0; \quad y_3 = -\frac{3}{2}\sqrt[3]{2} = -1{,}89$$

и, очевидно, невозможно переупорядочить y_1, y_2, y_3 так, чтобы получить $|y_\nu - x_\nu| \leqslant \varepsilon = 1$ $(\nu = 1, 2, 3)$.

ПРИЛОЖЕНИЕ II

ОТНОСИТЕЛЬНАЯ НЕПРЕРЫВНОСТЬ КОРНЕЙ АЛГЕБРАИЧЕСКИХ УРАВНЕНИЙ

1. В приложении I проблема непрерывности рассматривалась с точки зрения абсолютных погрешностей. Но при вычислениях часто приходится иметь дело с числами столь различных порядков, что невозможно обойтись без ограничений на *относительные погрешности*. Так бывает, в частности, в тех случаях, когда величины, участвующие в вычислениях, заданы с одинаковым числом значащих цифр. Нас интересует вопрос, можно ли гарантировать какое-либо число значащих цифр корней алгебраического уравнения, если коэффициенты этого уравнения заданы с определенным числом значащих цифр, независимо от их величины. Положительный ответ на этот вопрос дает следующая теорема.

2. Теорема. Пусть даны два многочлена:

$$p(z) = a_0 z^n + a_1 z^{n-1} + \ldots + a_n,$$
$$q(z) = b_0 z^n + b_1 z^{n-1} + \ldots + b_n, \qquad \text{(II. 1)}$$

где $a_0 a_n \neq 0$, т. е. $p(z)$ имеет n конечных корней x_ν $(\nu = 1, \ldots, n)$, отличных от нуля.

Если при некотором положительном числе τ, удовлетворяющем условию

$$4n\tau^{1/n} \leqslant 1, \qquad \text{(II. 2)}$$

имеют место неравенства

$$|b_\nu - a_\nu| \leqslant \tau |a_\nu| \quad (\nu = 0, 1, \ldots, n), \qquad \text{(II. 3)}$$

то n корней $y_1, \ldots, y_n$ многочлена $q(z)$ можно перенумеровать так, чтобы

$$\left| \frac{y_\nu - x_\nu}{x_\nu} \right| < 8n\tau^{1/n} \quad (\nu = 1, \ldots, n). \qquad \text{(II. 4)}$$

3. Доказательство. Положим $|x_\nu| = r_\nu$ $(\nu = 1, \ldots, n)$. Мы можем считать, что

$$r_1 \geqslant r_2 \geqslant \ldots \geqslant r_n; \quad |y_1| \geqslant |y_2| \geqslant \ldots \geqslant |y_n|.$$

Рассмотрим многочлен

$$P(z) = |a_0| \prod_{\nu=1}^{n} (z + r_\nu) = A_0 z^n + \ldots + A_n. \qquad \text{(II. 5)}$$

Так как

$$A_0 = |a_0|; \quad A_\nu \geqslant |a_\nu| \quad (\nu = 1, \ldots n), \qquad \text{(II. 6)}$$

то из (II.3) следует оценка

$$|b_\nu - a_\nu| \leqslant \tau A_\nu \quad (\nu = 1, \ldots, n). \qquad \text{(II. 7)}$$

В дальнейшем проведении доказательства мы используем оценку (II. 7) вместо (II. 3).

4. Пусть y_0 — любой корень многочлена $q(z)$. Обозначим через x_0 один из корней x_ν, для которого

$$\min_\nu \left| \frac{y_0 - x_\nu}{x_\nu} \right| = \left| \frac{y_0 - x_0}{x_0} \right|. \qquad \text{(II. 8)}$$

Из (II.7) следует, что

$$|p(y_0)| = |p(y_0) - q(y_0)| \leqslant \tau P(|y_0|).$$

Сократив обе части этого неравенства на $|a_0| = A_0$ и обозначая $|y_0| = s$, получим

$$\prod_{\nu=1}^{n} |y_0 - x_\nu| \leqslant \tau \prod_{\nu=1}^{n} (s + r_\nu). \qquad \text{(II. 9)}$$

5. Возьмем теперь число d, удовлетворяющее условию

$$0 < d < \frac{1 - \tau^{1/n}}{1 + \tau^{1/n}}, \qquad \text{(II. 10)}$$

и, значит, условию

$$\left(\frac{1+d}{1-d}\right)^n \tau < 1, \qquad \text{(II. 11)}$$

и разобьем множество чисел $r_\nu = |x_\nu|$ на три класса.

К первому классу отнесем первые l чисел $(r_1, \ldots, r_l)$, которые $\geqslant s/d$. Если таких r_ν не найдется, то положим

$l=0$. Для любого r_λ из первого класса и соответствующего x_λ имеем, очевидно,

$$\left|\frac{s+r_\lambda}{y_0-x_\lambda}\right| \leqslant \frac{s+r_\lambda}{r_\lambda-s} = \frac{1+\frac{s}{r_\lambda}}{1-\frac{s}{r_\lambda}} \leqslant \frac{1+d}{1-d} \quad (\lambda=1, \ldots, l). \quad \text{(II. 12)}$$

Ко *второму классу* отнесем последние $n-k$ чисел $(r_{k+1}, \ldots, r_n)$, которые $\leqslant sd$. Если таких $r_\varkappa$ не найдется, то положим $k=n$. Для любого $r_\varkappa$ из второго класса и соответствующего $x_\varkappa$ имеем, очевидно,

$$\left|\frac{s+r_\varkappa}{y_0-x_\varkappa}\right| \leqslant \frac{s+r_\varkappa}{s-r_\varkappa} = \frac{1+\frac{r_\varkappa}{s}}{1-\frac{r_\varkappa}{s}} \leqslant \frac{1+d}{1-d} \quad (\varkappa=k+1, \ldots, n). \text{(II. 13)}$$

Наконец, к *третьему классу* отнесем все r_μ, лежащие *строго между* s/d и sd. Число таких величин r_μ есть $m=k-l \geqslant 0$. Покажем, что $m>0$. Если бы было $m=0$, то для всех r_ν были бы справедливы неравенства (II. 12) и (II. 13) и тогда из неравенства (II. 9) следовало бы

$$1 \leqslant \tau \prod_{\nu=1}^{n} \frac{s+r_\nu}{|y_0-x_\nu|} \leqslant \tau \left(\frac{1+d}{1-d}\right)^n,$$

что противоречит условию (II.11).

6. Разделим обе части неравенства (II.9) на произведение тех $|y_0-x_\nu|$, которым соответствуют r_ν из первых двух классов. В силу (II. 12) и (II. 13) получим

$$\prod_{\mu=l+1}^{k} |y_0-x_\mu| \leqslant \tau \prod_{\mu=l+1}^{k} (s+r_\mu) \left(\frac{1+d}{1-d}\right)^{n-m},$$

откуда

$$\prod_{\mu=l+1}^{k} \left|\frac{y_0-x_\mu}{x_\mu}\right| \leqslant \tau \prod_{\mu=l+1}^{k} \left(\frac{s+r_\mu}{r_\mu}\right) \left(\frac{1+d}{1-d}\right)^{n-m}. \quad \text{(II. 14)}$$

Но для любого r_μ из третьего класса имеем

$$\frac{s+r_\mu}{r_\mu}=1+\frac{s}{r_\mu}<1+\frac{1}{d}=\frac{1+d}{d}.$$

Поэтому из (II. 8) и (II. 14) получаем

$$\left|\frac{y_0-x_0}{x_0}\right|^m\leqslant\left(\frac{1-d}{d}\right)^m\left(\frac{1+d}{1-d}\right)^n\tau. \qquad \text{(II.15)}$$

Из (II. 11) и (II. 15) заключаем, что

$$\left|\frac{y_0-x_0}{x_0}\right|^m\leqslant\left(\frac{1-d}{d}\right)^m;\quad \left|\frac{y_0-x_0}{x_0}\right|\leqslant\frac{1-d}{d}.$$

7. Последнее неравенство справедливо при всех значениях d, удовлетворяющих условию (II. 10). Поэтому оно должно быть справедливо и для предельного значения $\frac{1-\tau^{1/n}}{1+\tau^{1/n}}$, что дает

$$\left|\frac{y_0-x_0}{x_0}\right|\leqslant\frac{2\tau^{1/n}}{1-\tau^{1/n}}. \qquad \text{(II. 16)}$$

Правая часть этого неравенства, в силу (II. 2), меньше 1. Возьмем теперь любое число ε, удовлетворяющее условию

$$\frac{2\tau^{1/n}}{1-\tau^{1/n}}<\varepsilon<1, \qquad \text{(II. 17)}$$

и положим

$$\varepsilon r_\nu=\varepsilon_\nu. \qquad \text{(II. 18)}$$

Тогда из (II. 16) следует, что $|y_0-x_0|<\varepsilon r_0=\varepsilon_0$. Таким образом, каждый корень y_μ многочлена $q(z)$ лежит в *открытой* ε_ν-окрестности U_ν одного из корней x_ν, где U_ν есть внутренность круга с центром x_ν и радиусом ε_ν ($\nu=1,\ldots,n$).

8. Последнюю часть доказательства проведем аналогично тому, как это было сделано в приложении I. Назовем два корня x_ν, x_μ *соседними*, если

$$|x_\nu-x_\mu|\leqslant\varepsilon_\nu+\varepsilon_\mu. \qquad \text{(II. 19)}$$

Из (II. 18) получаем при этом

$$x_\mu = \frac{1+\vartheta\varepsilon}{1+\vartheta'\varepsilon} x_\nu, \quad |\vartheta|, |\vartheta'| \leqslant 1. \qquad \text{(II. 20)}$$

Назовем два корня x_{ν_1}, x_{ν_l} *связанными*, если можно указать такую последовательность $x_{\nu_1}, x_{\nu_2}, \ldots, x_{\nu_l}$ корней многочлена $p(z)$, в которой все рядом стоящие корни являются соседними. Без ограничения общности можно считать, что $l \leqslant n$. Повторно применяя (II. 20), получаем

$$x_{\nu_l} = x_{\nu_1} \prod_{\lambda=1}^{l} \frac{1+\vartheta_\lambda\varepsilon}{1+\vartheta'_\lambda\varepsilon}, \quad |\vartheta_\lambda|, |\vartheta'_\lambda| \leqslant 1, \qquad \text{(II. 21)}$$

что можно записать в виде

$$x_{\nu_l} = x_{\nu_1} \prod_{\lambda=1}^{n-1} \frac{1+\vartheta_\lambda\varepsilon}{1+\vartheta'_\lambda\varepsilon}, \quad |\vartheta_\lambda|, |\vartheta'_\lambda| \leqslant 1, \qquad \text{(II. 22)}$$

полагая ϑ_λ и ϑ'_λ равными 0 при $\lambda \geqslant l$.

9. Разобьем все x_ν на группы $g_1 \ldots, g_k$ таким образом, чтобы каждые два x_ν из одной группы были *связаны* и никакие два x_ν из разных групп не могли быть связаны. Для каждой группы построим объединение ε_ν - окрестностей U_ν корней x_ν, $G_\varkappa = \sum_{x_\nu \in g_\varkappa} U_\nu$. Граница $B_\varkappa$ любой области $G_\varkappa$ состоит из конечного числа дуг окружностей, причем каждая кривая $B_\varkappa$ отстоит на *положительное* расстояние от всех G_λ при $\lambda \neq \varkappa$.

Из (II. 16) и (II. 17) следует, что каждый корень y_ν лежит в одной из открытых областей $G_\varkappa$ и поэтому многочлен $q(z)$ не обращается в нуль ни на какой кривой $B_\varkappa$.

10. Докажем теперь, что в каждой области $G_\varkappa$ содержится столько же точек y_ν, сколько и x_ν. Действительно. если заменить многочлен $q(z)$ на многочлен

$$p(z) + t[q(z) - p(z)], \quad (0 \leqslant t \leqslant 1), \qquad \text{(II. 23)}$$

то условие (II. 7) будет выполнено с заменой τ на $t\tau$. Поэтому ни один многочлен вида (II. 23) не обращается в нуль на $B_\varkappa$. По лемме из п. 5 приложения I число

корней многочлена (II. 23) в области $G_\varkappa$ не зависит от t при $0 \leqslant t \leqslant 1$. Но этот многочлен совпадает с $p(z)$ при $t = 0$ и с $q(z)$ при $t = 1$. Следовательно, число точек x_ν, лежащих в области $G_\varkappa$, равно числу точек y_ν, лежащих там же.

11. Перенумеруем y_ν так, чтобы все y_ν, лежащие в области $G_\varkappa$, имели те же индексы, что и x_ν, лежащие в этой области. Конечно, это можно осуществить многими способами, если $G_\varkappa$ содержит больше чем один корень x_ν.

Рассмотрим какой-нибудь корень y_ν, лежащий в $G_\varkappa$. Из (II. 16) следует, что для некоторого x_μ из $G_\varkappa$ имеет место оценка

$$\left|\frac{y_\nu - x_\mu}{x_\mu}\right| < \varepsilon; \quad y_\nu = x_\mu(1 + \vartheta\varepsilon), \quad |\vartheta| \leqslant 1.$$

С другой стороны, корень x_ν связан с x_μ. Поэтому на основании (II. 22), получаем

$$y_\nu = x_\nu(1 + \vartheta\varepsilon)\prod_{\lambda=1}^{n-1}\frac{1 + \vartheta_\lambda\varepsilon}{1 + \vartheta'_\lambda\varepsilon},$$

$$\frac{y_\nu}{x_\nu} - 1 = (1 + \vartheta\varepsilon)\prod_{\lambda=1}^{n-1}\frac{1 + \vartheta_\lambda\varepsilon}{1 + \vartheta'_\lambda\varepsilon} - 1.$$

Отсюда находим следующую оценку

$$\left|\frac{y_\nu - x_\nu}{x_\nu}\right| \leqslant \varphi(\varepsilon), \quad \text{где} \quad \varphi(\varepsilon) = \frac{(1+\varepsilon)^n}{(1-\varepsilon)^{n-1}} - 1. \qquad \text{(II. 24)}$$

12. Нам остается доказать, что можно так выбрать число ε между числами

$$\delta = \frac{2\tau^{1/n}}{1 - \tau^{1/n}}$$

и 1, чтобы выполнялось неравенство $\varphi(\varepsilon) < 8n\tau^{1/n}$. Обозначим $\tau^{1/n} = \tau'$ и исследуем величину

$$\varphi(\delta) = \frac{(1+\tau')^n}{(1-\tau')(1-3\tau')^{n-1}} - 1 = \tau'\psi(\tau').$$

Так как по (II. 2) $\tau' \leqslant {}^1/_{4n}$, то $\psi(\tau')$ разлагается в степенной ряд с *положительными* коэффициентами и поэтому

$$\psi(\tau') \leqslant \psi\left(\frac{1}{4n}\right); \quad \varphi(\delta) \leqslant \tau'\psi\left(\frac{1}{4n}\right).$$

С другой стороны,

$$\psi\left(\frac{1}{4n}\right) = 4n\left[\frac{\left(1+\frac{1}{4n}\right)^n}{\left(1-\frac{3}{4n}\right)^{n-1}\left(1-\frac{1}{4n}\right)} - 1\right]. \quad \text{(II. 25)}$$

Первый член в квадратной скобке оценим так:

$$\frac{\left(1+\frac{1}{4n}\right)^n}{\left(1-\frac{3}{4n}\right)^{n-1}\left(1-\frac{1}{4n}\right)} = \frac{4n}{4n-1}\left(1+\frac{1}{4n}\right)^n \left(1+\frac{3}{4n-3}\right)^{n-1} <$$

$$< \left(1+\frac{1}{4n-1}\right)\left(1+\frac{\frac{1}{4}}{n}\right)^n\left(1+\frac{\frac{3}{4}}{n-1}\right)^{n-1} <$$

$$< \left(1+\frac{1}{4n-1}\right)e^{1/4}\,e^{3/4},$$

ибо при любом положительном целом n и любом положительном x:

$$e^x > \left(1+\frac{x}{n}\right)^n.$$

Таким образом, при $n \geqslant 3$ имеем

$$\psi\left(\frac{1}{4n}\right) < 4n\left(\frac{12}{11}e - 1\right) < 4n \cdot 1{,}9654.$$

При $n=2$ непосредственно вычисляем

$$\psi\left(\frac{1}{4n}\right) = 4n\left[\frac{\left(1+\frac{1}{8}\right)^2}{\left(1-\frac{3}{8}\right)\left(1-\frac{1}{8}\right)} - 1\right] = 4n \cdot \frac{46}{35} < 4n \cdot 1{,}9.$$

Итак, получаем окончательную оценку:

$$\psi(\delta) \leqslant 7{,}9\,n; \quad \varphi(\delta) \leqslant 7{,}9n\tau'. \quad \text{(II.26)}$$

Так как функция $\varphi(\varepsilon)$ непрерывна при $\varepsilon < 1$, то можно с самого начала выбрать ε так близко к δ, чтобы выполнялось неравенство

$$\varphi(\varepsilon) < 8n\tau',$$

что и доказывает оценку (II. 4).

13. Заметим еще, что если в левой части оценки (II. 4) заменить x_ν, y_ν на $|x_\nu|$, $|y_\nu|$, то правую часть можно заменить меньшей величиной, эквивалентной $2n\tau^{1/n}$ при $\tau \downarrow 0$[1]).

14. Одним из условий нашей теоремы является неравенство (II. 3), накладывающее очень жесткие требования на коэффициенты b_ν. В частности, при $a_\nu = 0$ должно быть также и $b_\nu = a_\nu = 0$; при очень малой величине a_ν величина b_ν также должна быть очень мала. Поэтому для приложений очень важно, что теорема остается справедливой, если условие (II. 3) заменить на (II. 7). Рассмотрим, например, уравнение $x^n - 1 = 0$. По условию (II. 3) получаем $|b_0 - 1| \leqslant \tau$, $|b_n + 1| \leqslant \tau$, $b_\nu = 0$ $(0 < \nu < n)$. С другой стороны, в нашем случае $A_\nu = \binom{n}{\nu}$ и, применяя (II.7), получаем существенно менее сильные условия на b_ν:

$$|b_0 - 1| \leqslant \tau, \ |b_n + 1| \leqslant \tau, \ |b_\nu| \leqslant \binom{n}{\nu} \tau \ (0 < \nu < n).$$

Однако для такого применения неравенства (II. 7) необходимо знать все $|x_\nu|$, а мы обычно их не знаем:

15. Мы переходим теперь к проблеме нахождения таких *положительных* величин T_ν, которые могут быть определены непосредственно по величинам $|a_\nu|$ и удовлетворяют неравенствам

$$|a_\nu| \leqslant T_\nu \leqslant A_\nu \quad (\nu = 1, \ldots, n-1). \tag{II.27}$$

Очевидно, что если такие величины найдены, то (II. 3) можно заменить условием

$$|b_\nu - a_\nu| \leqslant \tau T_\nu. \tag{II. 28}$$

---

[1]) Сравни с работой Ostrowski A., Mathematische Miszellen XXIV, Zur relativen Stetigkeit von Wurzeln algebraischer Gleichungen, *Jahresber. deut. Math. Ver.*, **58** (1956), 98—102.

Для нахождения таких величин мы воспользуемся неравенством, которое было известно еще Ньютону.

Если у многочлена с вещественными коэффициентами

$$\varphi(z) = c_0 z^n + \binom{n}{1} c_1 z^{n-1} + \ldots + \binom{n}{\nu} c_\nu z^{n-\nu} + \ldots + c_n,$$
$$c_0 \neq 0, \quad \text{(II.29)}$$

все n корней — вещественные числа, то

$$c_\nu^2 \geqslant c_{\nu-1} c_{\nu+1} \quad (\nu = 1, \ldots, n-1). \qquad \text{(II. 30)}$$

Это утверждение, очевидно, справедливо при $n = 2$, а в общем случае может быть доказано по индукции с использованием теоремы Ролля [1]).

16. В дальнейшем мы предполагаем, что

$$a_0 = 1. \qquad \text{(II. 31)}$$

Запишем многочлен $P(z)$ (II. 5) в виде

$$P(z) = \sum_{\nu=0}^{n} \binom{n}{\nu} c_\nu z^{n-\nu}, \quad \binom{n}{\nu} c_\nu = A_\nu, \quad c_0 = A_0 = 1, \quad \text{(II. 32)}$$

и положим

$$|a_\nu| = \binom{n}{\nu} k_\nu \quad (\nu = 0, 1, \ldots, n). \qquad \text{(II. 33)}$$

Из неравенств (II. 30) и (II. 6) следует, что

$$\sqrt{k_{\nu-1} k_{\nu+1}} \leqslant c_\nu \quad (\nu = 1, \ldots, n-1),$$

т. е.

$$\sqrt{\left(1+\frac{1}{\nu}\right)\left(1+\frac{1}{n-\nu}\right)|a_{\nu-1} a_{\nu+1}|} \leqslant A_\nu. \qquad \text{(II. 34)}$$

В частности, это неравенство верно при тех ν, при которых $a_\nu = 0$, $a_{\nu-1} a_{\nu+1} \neq 0$.

Поэтому в правой части неравенства (II. 3) можно заменить соответствующую величину $|a_\nu|$ на

$$\sqrt{\left(1+\frac{1}{\nu}\right)\left(1+\frac{1}{n-\nu}\right)|a_{\nu-1} a_{\nu+1}|}.$$

---

[1]) См., например, Hardy G. H., Littlewood J. E., Polya G., Inequalities, pp. 53, 54, Cambridge, Univ. Press, 1934 (русский перевод: Харди Г. Г., Литтльвуд Д. Е. и Полиа Г., Неравенства, М., ИЛ, 1948, стр. 69, 129).

Если подряд несколько чисел a_ν обращаются в нуль, точнее:

$$a_{\nu_1} \neq 0,\ a_{\nu_1+1} = a_{\nu_1+2} = \ldots = a_{\nu_2-1} = 0,\ a_{\nu_2} \neq 0, \qquad \text{(II. 35)}$$

то, как легко показать, соответствующие $|a_\nu|$ можно заменить в правой части (II. 3) на

$$\binom{n}{\nu}\binom{n}{\nu_2}^{\frac{\nu_1-\nu}{\nu_2-\nu_1}}\binom{n}{\nu_1}^{\frac{\nu-\nu_2}{\nu_2-\nu_1}} a_{\nu_2}^{\frac{\nu-\nu_1}{\nu_2-\nu_1}}\, a_{\nu_1}^{\frac{\nu_2-\nu}{\nu_2-\nu_1}}. \qquad \text{(II. 36)}$$

Мы дадим для этого случая геометрический способ построения величин T_ν.

17. Положим

$$\gamma_\nu = \ln c_\nu;\quad \alpha_\nu = \ln k_\nu\quad (\nu = 0, 1, \ldots, n); \qquad \text{(II. 37)}$$

при $k_\nu = 0$ мы считаем $\alpha_\nu = -\infty$.

Из (II. 30) и (II. 6) имеем

$$\gamma_\nu \geqslant \frac{1}{2}(\gamma_{\nu-1} + \gamma_{\nu+1}), \qquad \text{(II. 38)}$$

$$\alpha_\nu \leqslant \gamma_\nu. \qquad \text{(II.39)}$$

Неравенство (II. 38) допускает простую геометрическую интерпретацию. Если мы нанесем на плоскости (ν, γ) точки

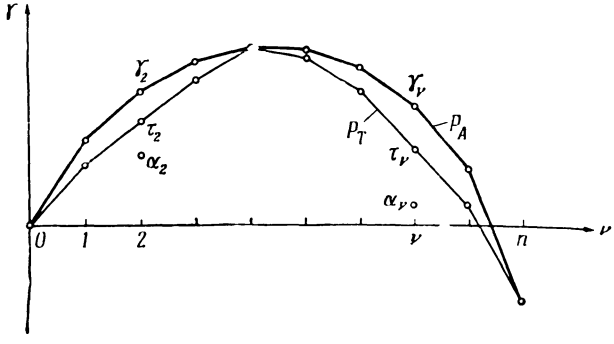

Рис. 7.

с координатами (ν, γ_ν), $\nu = 0, 1, \ldots, n$, и соединим их прямолинейными отрезками (см. рис. 7), то получим выпуклую ломаную линию P_A, обращенную *выпуклостью* вверх. Нанесем на том же рисунке точки (ν, α_ν). В силу неравенства (II.39) эти точки лежат *на ломаной P_A или ниже ее.*

Построим теперь ломаную линию P_T, соединяющую точки $(0, \alpha_0)$ и (n, α_n), так, чтобы она была *выпуклой вверх* и ее вершины находились только в точках (ν, α_ν), а остальные точки (ν, α_ν) лежали на ломаной P_T или ниже ее. Очевидно, что ни одна точка, принадлежащая P_T, не лежит выше P_A. Пусть для каждого ν величина τ_ν обозначает ординату точки, лежащей на P_T и имеющей абсциссу ν. При этом, очевидно,

$$\alpha_\nu \leqslant \tau_\nu \leqslant \gamma_\nu, \tag{II. 40}$$

и, следовательно, величины

$$T_\nu = \binom{n}{\nu} e^{\tau_\nu} \tag{II. 41}$$

удовлетворяют неравенству (II. 27). Таким образом, условие (II. 3) можно заменить условием (II. 28), в котором величины T_ν получены путем построения ломаной линии P_T.

ПРИЛОЖЕНИЕ III

ЯВНАЯ ФОРМУЛА ДЛЯ n-Й ПРОИЗВОДНОЙ ОБРАТНОЙ ФУНКЦИИ

1. Предположим, что $y = f(x)$ дифференцируема достаточное число раз, и обозначим $f^{(\nu)}(x)$ через y_ν. Пусть $x = \varphi(y)$ — функция, обратная к $y = f(x)$. Тогда, как показано в гл. 2,

$$\varphi^{(n)}(y) = y_1^{-(2n-1)} X_n(y_1, \ldots, y_n), \qquad \text{(III. 1)}$$

где X_n — многочлен относительно $y_1, \ldots, y_n$:

$$X_n = \sum a_{\alpha_1 \ldots \alpha_n} y_1^{\alpha_1} \ldots y_n^{\alpha_n}, \qquad \text{(III. 2)}$$

а показатели $\alpha_1, \ldots, \alpha_n$ удовлетворяют условиям

$$\alpha_\nu \geqslant 0, \quad \sum_{\nu=1}^{n} \alpha_\nu = n - 1, \quad \sum_{\nu=1}^{n} \nu\alpha_\nu = 2n - 2. \qquad \text{(III.3)}$$

Мы покажем, что каждый одночлен, удовлетворяющий условиям (III. 3), входит в X_n с ненулевым коэффициентом, причем этот коэффициент определяется по формуле

$$a_{\alpha_1 \ldots \alpha_n} = (-1)^{n-\alpha_1-1} \frac{(2n-\alpha_1-2)!}{\alpha_2!\,(2!)^{\alpha_2}\,\alpha_3!\,(3!)^{\alpha_3} \ldots \alpha_n!\,(n!)^{\alpha_n}}. \qquad \text{(III.4)}$$

2. Рассмотрим сначала член, содержащий y_n. Из условий (III. 3) следует, что при $n > 1$

$$\alpha_n = 1, \quad \alpha_1 = n - 2,$$

а все остальные показатели α_ν $(1 < \nu < n)$ равны нулю. Поэтому коэффициент (III. 4) равен -1 в соответствии с формулой (2.12). Итак, наше утверждение справедливо при $\alpha_n > 0$ и $n > 1$.

Мы будем доказывать не формулу (III. 4), а соответствующую формулу для коэффициентов производной $\varphi^{(n)}(y)$.

Доказательство проведем по индукции. Разделим (III. 2) на y_1^{2n-1} и положим

$$2n - \alpha_1 - 1 = \beta_1, \quad \alpha_\nu = \beta_\nu \quad (\nu > 1).$$

Тогда из (III. 2) и (III. 4) получим выражение

$$\varphi^{(n)}(y) = \sum (-1)^{n+\beta_1} \frac{(\beta_1 - 1)!}{\beta_2!\,(2!)^{\beta_2} \ldots \beta_n!\,(n!)^{\beta_n}} y_1^{-\beta_1} y_2^{\beta_2} \ldots y_n^{\beta_n}, \tag{III. 5}$$

где β_ν подчинены условиям

$$\beta_1 \leqslant 2n - 1, \quad \beta_\nu \geqslant 0 \quad (\nu > 1), \quad \sum_{\nu > 1} \beta_\nu = \beta_1 - n, \tag{III. 6}$$

$$\sum_{\nu > 1} \nu\beta_\nu = \beta_1 - 1. \tag{III.7}$$

3. Мы уже знаем, что при $n > 1$ наши утверждения справедливы для члена суммы (III. 5), у которого $\beta_n > 0$. При $n = 1$ условия (III. 6) и (III. 7) дают $\beta_1 = n = 1$, и мы получаем из (III. 5) известную формулу $\varphi'(y) = \frac{1}{y_1}$.

При $n = 2$ условия (III. 6) и (III. 7) приводят к системе двух уравнений $\beta_2 = \beta_1 - 2$, $2\beta_2 = \beta_1 - 1$, единственным решением которой является

$$\beta_1 = 3; \quad \beta_2 = 1.$$

В соответствии с (III. 5) мы получаем в этом случае $\varphi''(y) = -\frac{y_2}{y_1^3}$, что также верно.

4. Предположим теперь, что формула (III. 5) дает правильное значение $\varphi^{(n)}$ при некотором $n \geqslant 2$, и докажем соответствующую формулу для $\varphi^{(n+1)}$:

$$\varphi^{(n+1)}(y) = \sum T_{\gamma_1 \ldots \gamma_{n+1}} y_1^{-\gamma_1} y_2^{\gamma_2} \ldots y_{n+1}^{\gamma_{n+1}}, \tag{III. 8}$$

где

$$T_{\gamma_1 \ldots \gamma_{n+1}} = \\ = (-1)^{n+1+\gamma_1} \frac{(\gamma_1 - 1)!}{\gamma_2!\,(2!)^{\gamma_2} \ldots \gamma_n!\,(n!)^{\gamma_n}\,\gamma_{n+1}!\,[(n+1)!]^{\gamma_{n+1}}}. \tag{III. 9}$$

При доказательстве (III. 9) мы предположим, что $\gamma_{n+1}=0$, так как справедливость формулы (III.9) при $\gamma_{n+1}>0$ следует из п. 2.

Показатели γ_ν удовлетворяют условиям, аналогичным (III. 6) и (III. 7), причем условие, аналогичное (III. 7), имеет вид

$$\sum_{\nu\geqslant 2} \nu\gamma_\nu = \gamma_1 - 1. \qquad \text{(III. 10)}$$

Нетрудно показать, что $\gamma_1 - 1 > 0$. Действительно, в противном случае из (III. 10) получаем

$$\gamma_1 = 1, \quad \gamma_2 = \gamma_3 = \ldots = \gamma_n = \gamma_{n+1} = 0.$$

Но из условия, аналогичного (III. 6), следует, что

$$\gamma_2 + \ldots + \gamma_n + \gamma_{n+1} = \gamma_1 - (n+1).$$

Поэтому $n+1=1$, что противоречит условию $n>1$.

5. Член суммы (III. 8), содержащий $y_1^{-\gamma_1} y_2^{\gamma_2} \ldots y_{n+1}^{\gamma_{n+1}}$, получается из различных членов суммы (III. 5) при следующем процессе дифференцирования

$$\varphi^{(n+1)}(y) = \frac{1}{y_1}\left(y_2 \frac{\partial}{\partial y_1} + y_3 \frac{\partial}{\partial y_2} + \ldots + y_n \frac{\partial}{\partial y_{n-1}} + \right.$$
$$\left. + y_{n+1} \frac{\partial}{\partial y_n}\right) \varphi^{(n)}(y); \qquad \text{(III. 11)}$$

мы хотим определить «вклады» различных членов формулы (III. 11) в $T_{\gamma_1 \ldots \gamma_{n+1}}$. Рассмотрим прежде всего «вклад» члена $\frac{y_2}{y_1}\frac{\partial}{\partial y_1}$. Очевидно, что этот оператор должен быть применен к члену суммы (III. 5), в котором

$$\beta_1 = \gamma_1 - 2, \quad \beta_2 = \gamma_2 - 1, \quad \beta_\mu = \gamma_\mu \quad (\mu = 3, \ldots, n).$$

Но тогда «вклад» этого члена в $T_{\gamma_1 \ldots \gamma_{n+1}}$ равен

$$-(\gamma_1 - 2)(-1)^{n+\gamma_1-2} \frac{(\gamma_1-3)!}{(\gamma_2-1)!\,(2!)^{\gamma_2-1}\,(\gamma_3)!\,(3!)^{\gamma_3}\ldots\gamma_n!\,(n!)^n} =$$
$$= \frac{2\gamma_2}{\gamma_1-1} T_{\gamma_1 \ldots \gamma_{n+1}}.$$

6. Рассмотрим теперь «вклад» оператора $\frac{y_{\nu+1}}{y_1}\frac{\partial}{\partial y_\nu}$ при $\nu > 1$. Этот оператор должен быть применен к члену суммы (III. 5), в котором

$$\beta_1 = \gamma_1 - 1, \quad \beta_\nu = \gamma_\nu + 1; \quad \beta_{\nu+1} = \gamma_{\nu+1} - 1,$$
$$\beta_\mu = \gamma_\mu \quad (\mu \neq 1, \nu, \nu+1).$$

Поэтому «вклад» данного оператора в $T_{\gamma_1 \dots \gamma_{n+1}}$ есть

$$\frac{(\gamma_\nu + 1)(-1)^{n+\gamma_1-1}(\gamma_1 - 2)!}{\dots(\gamma_\nu+1)!(\nu!)^{\gamma_\nu+1}(\gamma_{\nu+1}-1)![(\nu+1)!]^{\gamma_{\nu+1}-1}\dots}, \quad \text{(III. 12)}$$

причем невыписанные в явном виде множители знаменателя, соответствующие $\gamma_2, \dots, \gamma_{\nu-1}, \gamma_{\nu+2}, \dots, \gamma_n$, совпадают с соответствующими множителями формулы (III. 9). Очевидно, что дробь (III. 12) равна

$$\frac{(\nu+1)!\,\gamma_{\nu+1}}{\nu!\,(\gamma_1 - 1)} T_{\gamma_1 \dots \gamma_{n+1}} = \frac{(\nu+1)\gamma_{\nu+1}}{\gamma_1 - 1} T_{\gamma_1 \dots \gamma_{n+1}}.$$

Суммируя все «вклады», получаем

$$\frac{T_{\gamma_1 \dots \gamma_{n+1}}}{\gamma_1 - 1}\left(2\gamma_2 + \sum_{\nu=2}^{n} (\nu+1)\gamma_{\nu+1}\right). \quad \text{(III. 13)}$$

Но в силу формулы (III. 10) выражение, заключенное в скобки, равно $\gamma_1 - 1$; поэтому сумма всех «вкладов» равна $T_{\gamma_1 \dots \gamma_{n+1}}$, что и доказывает формулу (III. 8).

7. *Частные случаи.* Если $f(x)$ — *квадратичный многочлен*, то $y_3 = y_4 = \dots = 0$ и в формуле (III. 5) имеем $\beta_2 = n - 1$, $\beta_1 = 2n - 1$,

$$\varphi^{(n)}(y) = (-1)^{n-1}\, 3\cdot 5 \dots (2n-3)\frac{y_2^{n-1}}{y_1^{2n-1}}. \quad \text{(III.14)}$$

Если $f(x)$ — *многочлен третьей степени*, то

$$y_4 = y_5 = \dots = 0$$

и

$$\varphi^{(n)}(y) = \sum (-1)^{n+\beta_1}\frac{(\beta_1 - 1)!}{\beta_2!\,2^{\beta_2}\,\beta_3!\,6^{\beta_3}} y_1^{-\beta_1} y_2^{\beta_2} y_3^{\beta_3},$$

где

$$\beta_2 \geqslant 0; \quad \beta_3 \geqslant 0; \quad \beta_1 \leqslant 2n-1; \quad \beta_2+\beta_3=\beta_1-n;$$
$$2\beta_2+3\beta_3=\beta_1-1.$$

Полагая $\beta_3=\mu$, получаем

$$\beta_2=n-2\mu-1, \quad \beta_1=2n-\mu-1.$$

При этом $\mu \geqslant 0$, но $\mu \leqslant \frac{n-1}{2}$, так как $\beta_2 \geqslant 0$. Окончательно получаем в этом случае следующее выражение для производной:

$$\varphi^{(n)}(y)=y_2^{n-1}\,y_1^{1-2n}\times$$

$$\times\sum_{\mu=0}^{\left[\frac{n-1}{2}\right]}(-1)^{n-\mu-1}\frac{(2n-\mu-2)!}{(n-2\mu-1)!\,\mu!\,2^{n-\mu-1}\,3^{\mu}}\cdot\left(\frac{y_1y_3}{y_2^2}\right)^{\mu}. \quad \text{(III. 15)}$$

Здесь $\left[\frac{n-1}{2}\right]$ обозначает целую часть числа $\frac{n-1}{2}$.

ПРИЛОЖЕНИЕ IV

АНАЛОГ МЕТОДА СЕКУЩИХ ДЛЯ СЛУЧАЯ ДВУХ УРАВНЕНИЙ С ДВУМЯ НЕИЗВЕСТНЫМИ

1. Пусть требуется приближенно решить систему уравнений

$$F(P) \equiv F(x, y) = 0, \quad G(P) \equiv G(x, y) = 0.$$

Предположим, что нам известны значения функций F и G в трех точках P_1, P_2, P_3:

$$F(P_\nu), \quad G(P_\nu) \qquad (\nu = 1, 2, 3)^{1)}.$$

Определим две линейные функции от x, y:

$$L_1 \equiv ax + by + c; \quad L_2 \equiv ex + fy + g,$$

удовлетворяющие шести условиям[2])

$$L_1(P_\nu) = F(P_\nu), \quad L_2(P_\nu) = G(P_\nu) \qquad (\nu = 1, 2, 3), \qquad \text{(IV. 1)}$$

и разрешим систему уравнений

$$L_1(x, y) = 0, \quad L_2(x, y) = 0 \qquad \text{(IV. 2)}$$

относительно x, y.

Для того, чтобы из системы (IV. 1), (IV. 2) исключить a, b, c, e, f, g, исключим сначала c и g, вычитая при каждом ν из уравнений системы (IV. 2) соответствующие уравнения системы (IV. 1). Тогда получим

$$\left.\begin{aligned} a(x - x_\nu) + b(y - y_\nu) &= -F(P_\nu), \\ e(x - x_\nu) + f(y - y_\nu) &= -G(P_\nu) \end{aligned}\right\} \quad (\nu = 1, 2, 3). \qquad \text{(IV. 3)}$$

---

[1]) Этот случай рассмотрен Гауссом в Theoria Motus Corporum Coelestium in Sectionibus Conicis Solem Ambientium, 1809, № 120, pp. 136—138; Gauss C. F., Werke, vol. VII, pp. 150—152.

[2]) Функции L_1, L_2 будут определены при любых правых частях уравнений (IV. 1), если определитель

$$\begin{vmatrix} 1 & 1 & 1 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix} = \begin{vmatrix} 1 \\ x_\nu \\ y_\nu \end{vmatrix} \qquad (\nu = 1, 2, 3)$$

отличен от нуля.—*Прим. перев.*

2. Из уравнений (IV. 3) следует, что матрица

$$\left(x - x_\nu,\ y - y_\nu,\ F(P_\nu),\ G(P_\nu)\right) \quad (\nu = 1, 2, 3)$$

имеет ранг не больше двух. Мы получаем отсюда два уравнения

$$\begin{vmatrix} x - x_\nu \\ F(P_\nu) \\ G(P_\nu) \end{vmatrix} = 0, \quad \begin{vmatrix} y - y_\nu \\ F(P_\nu) \\ G(P_\nu) \end{vmatrix} = 0, \quad (\nu = 1, 2, 3), \qquad \text{(IV. 4)}$$

которые были даны Гауссом для нахождения четвертого улучшенного приближения к решению системы по трем данным приближениям P_1, P_2, P_3.

Если определитель

$$\Delta = \begin{vmatrix} 1 \\ F(P_\nu) \\ G(P_\nu) \end{vmatrix} \quad (\nu = 1, 2, 3) \qquad \text{(IV. 5)}$$

отличен от нуля, то решение системы (IV. 4) дается формулами

$$x = \frac{1}{\Delta}\begin{vmatrix} x_\nu \\ F(P_\nu) \\ G(P_\nu) \end{vmatrix}, \quad y = \frac{1}{\Delta}\begin{vmatrix} y_\nu \\ F(P_\nu) \\ G(P_\nu) \end{vmatrix} \quad (\nu = 1, 2, 3). \qquad \text{(IV. 6)}$$

3. Условие $\Delta \neq 0$ эквивалентно требованию, чтобы три точки $\left(F(P_\nu),\ G(P_\nu)\right)$ $(\nu = 1, 2, 3)$ не лежали на одной прямой. С другой стороны, из (IV. 5) и (IV. 3) следует, что

$$\Delta = (af - be)\begin{vmatrix} 1 \\ x_\nu \\ y_\nu \end{vmatrix} \quad (\nu = 1, 2, 3). \qquad \text{(IV. 7)}$$

Итак, для того чтобы определитель Δ был отличен от нуля, необходимо, чтобы три точки P_1, P_2, P_3 не лежали на одной прямой. Гаусс предложил выбирать эти точки так, чтобы выполнялись два условия: $x_2 = x_1$, $y_3 = y_1$, т. е. чтобы точки P_1, P_2, P_3 лежали в вершинах прямоугольного треугольника.

После того как по (IV. 6) найдено четвертое приближение P_4, одна из точек P_1, P_2 или P_3 отбрасывается, а для оставшегося триплета повторяются те же вычисления.

Этот метод, по-видимому, до сих пор практически не был известен и почти не применялся в вычислительной практике, а исследование его сходимости, связанное с довольно тонкими рассуждениями, никогда не проводилось.

ПРИЛОЖЕНИЕ V

УЛУЧШЕНИЕ ИТЕРАЦИИ ПО МЕТОДУ СТЕФФЕНСЕНА

1. Для улучшения итерации, которую дает функция $\psi(x)$, Стеффенсен в 1933 году[1]) предложил следующий метод: начиная с x_0, вычислить значения $x_1 = \psi(x_0)$, $x_2 = \psi(x_1)$ и затем применить regula falsi к уравнению

$$F(x) \equiv x - \psi(x) = 0,$$

взяв за исходные точки x_0 и x_1. Таким образом, мы получаем новое приближение

$$y_1 = \frac{x_0 F(x_1) - x_1 F(x_0)}{F(x_1) - F(x_0)}$$

или, так как

$$F(x_0) = x_0 - x_1; \quad F(x_1) = x_1 - x_2,$$

приближение

$$y_1 = \frac{x_0 x_2 - x_1^2}{x_0 - 2x_1 + x_2}.$$

Положив теперь $y_0 = x_0$, получим окончательно итерацию $y_1 = \Psi(y_0)$, где

$$\Psi(y) = \frac{y\psi[\psi(y)] - \psi(y)^2}{y - 2\psi(y) + \psi[\psi(y)]}. \qquad \text{(V. 1)}$$

---

[1]) Steffensen J.F., Remarks on iteration, *Skand. Aktuar. Tidskr.*, 1933, 16, 64—72.

[В русской литературе этот метод известен как «метод Эйткена». Эйткен в 1931 г. предложил итерационный процесс, основанный на формуле

$$x_3 = x_2 - (\Delta x_1)^2 / \Delta^2 x_0$$

и совпадающий по существу с итерацией Стеффенсена.—*Прим. перев.*]

2. Стеффенсен показал на примерах, что эта итерация сходится очень быстро к корню функции $F(x)$ даже тогда, когда $|\psi'| > 1$. Виллерс в своем учебнике ([7], стр. 259) утверждает, что этот метод «всегда работает»; в статье в Zeitschrift für angewandte Mathematik und Mechanik[1]) он дал изящную геометрическую иллюстрацию применения метода Стеффенсена. Хаусхолдер ([1], стр. 126—128) показал, что при некоторых специальных условиях неподвижная точка итерации, определяемой функцией $\psi(x)$, становится точкой притяжения для $\Psi(y)$. Мы докажем здесь четыре теоремы, показывающие, что при весьма общих условиях корень ζ уравнения $F(x) = 0$ является точкой притяжения для итерации с функцией[2]) $\Psi(y)$.

3. При доказательстве этих теорем мы будем предполагать, что $\zeta = 0$. Такое предположение не приводит к ограничению общности. Действительно, если $\zeta \neq 0$, то положим

$$z = x - \zeta; \quad F^*(z) = F(x) = F(z + \zeta), \quad \psi^*(z) = \psi(z + \zeta) - \zeta$$

и

$$\Psi^*(y) = \frac{y\psi^*[\psi^*(y)] - \psi^*(y)^2}{y - 2\psi^*(y) + \psi^*[\psi^*(y)]}.$$

При этом неподвижной точкой для $\psi^*(z)$ будет точка $z = 0$; итерации $x_\nu = \psi(x_{\nu-1})$, начинающейся с точки x_0, будет соответствовать итерация $z_\nu = \psi^*(z_{\nu-1}) = x_\nu - \zeta$, начинающаяся с точки $z_0 = x_0 - \zeta$; наконец, будет справедливо соотношение

$$\Psi^*(y) = \Psi(y + \zeta) - \zeta.$$

В дальнейшем мы предполагаем также, что существует производная $\psi'(\zeta)$; значение $\psi'(\zeta)$ обозначим через α.

Теорема 1. *Если* $\alpha \neq 0$, $\alpha \neq 1$, *то для итерирующей функции* $\Psi(x)$ *точка* ζ *всегда будет точкой притяжения, причем* $\Psi'(\zeta) = 0$.

---

[1]) Willers F. A., *Z. angew. Math. Mech.*, **22**, 125, 126 (1948).

[2]) Мы доопределяем функцию $\Psi(y)$ в точке ζ, полагая $\Psi(\zeta) = \zeta$. Если $\Psi'(\zeta) \neq 1$, то, как легко показать, $\lim\limits_{y \to \zeta} \Psi(y) = \zeta$. Таким образом, в этом случае функция $\Psi(y)$ будет непрерывна в точке ζ. В случае $\Psi'(\zeta) = 1$ требуются дополнительные предположения; см., например, теорему IV.—*Прим. перев.*

4. Доказательство. Предположим, без ограничения общности, что $\zeta = 0$. Тогда при $x \to 0$ имеем

$$\psi(x) = \alpha x + o(x),$$

$$\psi[\psi(x)] = \alpha[\alpha x + o(x)] + o[\alpha x + o(x)] = \alpha^2 x + o(x),$$

$$\psi[\psi(x)] - 2\psi(x) + x = (\alpha - 1)^2 x + o(x),$$

$$\psi(x)^2 = \alpha^2 x^2 + o(x^2),$$

$$x\psi[\psi(x)] - \psi(x)^2 = o(x^2).$$

Поэтому из (V. 1) следует, что $\Psi(x) = o(x)$, т. е. $\Psi'(0) = 0$. Теорема 1 доказана.

Результат теоремы 1 можно улучшить, зная порядок бесконечно малой $\psi(x) - \zeta - \alpha(x - \zeta)$ при $x \to \zeta$.

5. Теорема II. *Пусть для некоторого* $\lambda > 1$ *выражение*

$$E(x) = \frac{\psi(x) - \zeta - \alpha(x - \zeta)}{|x - \zeta|^\lambda} \tag{V. 2}$$

при $x \to \zeta$ *либо* (a) *остается ограниченным, т. е.* $E = O(1)$, *либо* (b) *стремится к нулю, т. е.* $E = o(1)$.

Если при этом $\alpha(\alpha - 1) \neq 0$, *то имеем соответственно*

$$\Psi(x) - \zeta = O(|x - \zeta|^\lambda), \tag{V. 3a}$$

или

$$\Psi(x) - \zeta = o(|x - \zeta|^\lambda). \tag{V. 3b}$$

Если $\alpha > 0$ *и одно из предположений* (a) *или* (b) *выполнено лишь при одностороннем стремлении* x к ζ, *то при том же одностороннем стремлении справедлива соответствующая оценка* (V. 3a) *или* (V. 3b).

6. Доказательство. Предположим, без потери общности, что $\zeta = 0$. При этом из (V. 2) имеем

$$\psi(x) = \alpha x + E(x)|x|^\lambda,$$

$$\psi[\psi(x)] = \alpha[\alpha x + E(x)|x|^\lambda] + E[\psi(x)]|\alpha x + E(x)|x|^\lambda|^\lambda =$$
$$= \alpha^2 x + \{\alpha E(x) + |\alpha|^\lambda E[\psi(x)]\}|x|^\lambda + o(|x|^\lambda).$$

Поэтому

$$\psi[\psi(x)] - 2\psi(x) + x = (\alpha - 1)^2 x + o(x), \tag{V. 4}$$

$$\psi(x)^2 = \alpha^2 x^2 + 2\alpha E(x)|x|^\lambda x + o(|x|^{\lambda+1}),$$

$$x\psi[\psi(x)] - \psi(x)^2 =$$
$$\{|\alpha|^\lambda E[\psi(x)] - \alpha E(x)\}|x|^\lambda x + o(|x|^{\lambda+1}). \tag{V. 5}$$

Деля (V. 5) на (V. 4), получаем

$$\Psi(x) = \frac{|x|^{\lambda}}{(\lambda-1)^2}\{|\alpha|^{\lambda} E[\psi(x)] - \alpha E(x)\} + o(|x|^{\lambda}). \quad (V.\ 6)$$

Заметим теперь, что $\psi(x) \to 0$ при $x \to 0$; поэтому в случаях (a) или (b) утверждение (V. 3a) или соответственно (V.3b) непосредственно следует из (V. 6).

7. Если $\alpha > 0$ и, например, $E(x) = O(1)$ при $x \downarrow 0$, то $\psi(x)$ положительна при достаточно малых $x > 0$, поэтому $E[\psi(x)]$ также будет $O(1)$. Но тогда (V. 3a) следует из (V. 6) при $x \downarrow 0$. Так же можно доказать утверждения, относящиеся к случаю $x \uparrow 0$. Теорема II доказана.

8. Если $\alpha = 0$, то теоремы I и II остаются справедливыми, но ничего не говорят об улучшении сходимости. Однако в этом случае теорема II может быть усилена.

Теорема III. *Пусть* $\alpha = 0$ *и для некоторого* $\lambda > 1$ *выражение*

$$E(x) = \frac{\psi(x) - \zeta}{|x - \zeta|^{\lambda}} \quad (V.\ 7)$$

при $x \to \zeta$ *либо* (a) *остается ограниченным, либо* (b) *стремится к нулю. Тогда справедливы соответственно или оценка*

$$\Psi(x) - \zeta = O(|x - \zeta|^{2\lambda-1}), \quad (V.\ 8a)$$

или оценка

$$\Psi(x) - \zeta = o(|x - \zeta|^{2\lambda-1})\ ^{1)}. \quad (V.\ 8b)$$

---

[1]) Хаусхолдер [1] получил результаты, соответствующие теоремам II и III, в предположении, что $\psi(x)$ может быть разложена по целым (по-видимому) степеням $x - \zeta$. Кроме того, Хаусхолдер дал обобщение метода Стеффенсена, показав, что по любым *двум* итерирующим функциям $\psi_1(x)$ и $\psi_2(x)$ можно построить новую итерирующую функцию $\Psi(x)$, дающую обычно более быструю сходимость, чем ψ_1 и ψ_2. В случае $\psi_1 = \psi_2$ формула Хаусхолдера совпадает с формулой Стеффенсена. С другой стороны, более подробное изучение метода Хаусхолдера показывает, что комбинация по его методу функций ψ_1 и ψ_2 может быть особенно полезна в том случае, когда ψ_2 есть комбинация ψ_1 и ψ_1. См. статьи автора: Über Verfahren von Steffensen und Householder zur Verbesserung der Konvergenz von Iterationen, *Z. angew. Math. Phys.*, 1956, 7, 218—219; On the Convergence of the Rayleigh Quotient Iteration for the Computation of Characteristic Roots and Vectors VI. (Usual Rayleigh Quotient for Nonlinear Elementary Divisors), *Archive for Rational Mechanics and Analysis*, 4 (1959), 153—165.

9. Доказательство. Предположим, без ограничения общности, что $\zeta = 0$. Тогда имеем

$$\psi(x) = E(x)\,|x|^{\lambda}, \quad \psi(x)^2 = E(x)^2\,|x|^{2\lambda},$$
$$\psi[\psi(x)] = E[\psi(x)]\,|E(x)\,|x|^{\lambda}|^{\lambda} = E[\psi(x)]\,|E(x)|^{\lambda}\,|x|^{\lambda^2}.$$

Так как $E[\psi(x)] = O(1)$, мы получаем отсюда

$$\psi[\psi(x)] - 2\psi(x) + x = x + O(|x|^{\lambda})$$

и, так как $\lambda^2 + 1 > 2\lambda$,

$$x\psi[\psi(x)] = O(|x|^{\lambda^2+1}) = o(|x|^{2\lambda}).$$

Далее,

$$x\psi[\psi(x)] - \psi(x)^2 = -E(x)^2\,|x|^{2\lambda} + o(|x|^{2\lambda}).$$

Наконец, из (V.1) получаем

$$\Psi(x) = -E(x)^2 x\,|x|^{2\lambda-2} + o(|x|^{2\lambda-1}), \qquad \text{(V. 9)}$$

откуда непосредственно следуют соотношения (V. 8a), (V. 8b).

10. Рассмотрим, наконец, случай $\psi'(\zeta) = \alpha = 1$.

Теорема IV. *Пусть $\alpha = 1$. Если $\psi'(x)$ непрерывна в окрестности точки ζ и если для некоторого $\lambda > 1$ справедливо соотношение*

$$\psi'(x) - 1 = T(x)\,|x - \zeta|^{\lambda-1}, \qquad \text{(V. 10)}$$

где $T(x)$ стремится к пределу $\gamma \neq 0$ либо при $x \uparrow \zeta$, либо при $x \downarrow \zeta$, то для соответствующей односторонней производной имеем

$$\Psi'(\zeta) = 1 - \frac{1}{\lambda} \qquad \text{(V. 11)}$$

и ζ является точкой притяжения (с соответствующей стороны) для итерирующей функции $\Psi(x)$.

11. Доказательство. Предположим, без потери общности, что $\zeta = 0$. Заменив в (V. 1) y на x, вычтем x из обеих частей полученного равенства:

$$\Psi(x) - x = \frac{-[\psi(x) - x]^2}{\psi[\psi(x)] - 2\psi(x) + x}.$$

Положим

$$g(x) = \psi(x) - x. \qquad \text{(V. 12)}$$

Тогда

$$\Psi(x) - x = -\frac{g(x)^2}{g[\psi(x)] - g(x)}. \qquad \text{(V. 13)}$$

Из теоремы о среднем значении следует, что

$$g[\psi(x)] - g(x) = [\psi(x) - x]\, g'(\xi) = g(x)\, g'(\xi), \qquad \text{(V. 14)}$$

где ξ лежит между x и $\psi(x)$, т. е.

$$\xi = x + \vartheta(\psi(x) - x) = x + \vartheta g(x), \qquad 0 < \vartheta < 1.$$

Из формулы (V. 10) при $\zeta = 0$ интегрированием получаем

$$g(x) = \psi(x) - x = \int_0^x T(x)\,|x|^{\lambda-1} dx = \frac{\gamma}{\lambda}\, x\,|x|^{\lambda-1} + o(x^\lambda).$$

Поэтому

$$\xi = x + o(x).$$

С другой стороны, из формул (V. 10) и (V. 12) имеем

$$g'(\xi) = \psi'(\xi) - 1 = T(\xi)\,|\xi|^{\lambda-1} = \gamma\,|x|^{\lambda-1} + o(|x|^{\lambda-1});$$

таким образом,

$$\Psi(x) - x = -\frac{g(x)}{g'(\xi)} \sim -\frac{x}{\lambda}. \qquad \text{(V. 15)}$$

Соотношение (V. 11) непосредственно следует из (V. 15).

13. Условие $\alpha = 1$, очевидно, эквивалентно условию $F'(\zeta) = 0$. В этом случае ζ обычно является кратным корнем уравнения $F(x) = 0$, и по теореме 5.2 итерация с $\psi(x)$ сходится крайне медленно, если сходится вообще. С другой стороны, из теоремы IV следует, что итерация с $\Psi(x)$ имеет линейную сходимость. Если мы теперь применим метод Стеффенсена к $\Psi(x)$, то в силу теоремы I получим даже сверхлинейную сходимость.

Если, как это обычно бывает, λ известно, то можно применить формулу (4.8). В этом случае мы получим сверхлинейную сходимость, даже не применяя метод Стеффенсена еще раз, а просто заменяя $\Psi(x)$ на

$$\Psi^*(x) = +\lambda\left[\Psi(x) - \left(1 - \frac{1}{\lambda}\right)x\right] = +\lambda\Psi(x) - (\lambda - 1)\,x.$$

Отметим, наконец, что в любом случае, если итерация с $\Psi(x)$ сходится к некоторому пределу ζ, то в силу формулы (V. 13) число ζ является корнем уравнения $g(x) = 0$, т. е. $\psi(x) - x = 0$.

ПРИЛОЖЕНИЕ VI

АЛГОРИТМ НЬЮТОНА В СЛУЧАЕ КВАДРАТИЧНЫХ МНОГОЧЛЕНОВ

1. Теорема. Если в квадратном уравнении

$$f(x) \equiv (x-\xi)(x-\eta) = 0 \qquad \text{(VI. 1)}$$

имеем

$$|\xi - x_0| < |\eta - x_0|, \qquad \text{(VI. 2)}$$

то алгоритм Ньютона с начальной точкой x_0 сходится к значению ξ. То же имеет место в случае $\xi = \eta$.

Если же

$$|\xi - x_0| = |\eta - x_0|, \quad \xi \neq \eta, \qquad \text{(VI. 3)}$$

то алгоритм Ньютона с начальной точкой x_0 расходится.

2. Доказательство. Положим

$$\xi - x_\nu = a_\nu, \quad \eta - x_\nu = b_\nu, \quad a_0 = a, \quad b_0 = b. \qquad \text{(VI. 4)}$$

По определению, имеем [1])

$$x_{\nu+1} = x_\nu - \frac{(x_\nu - \xi)(x_\nu - \eta)}{2x_\nu - \xi - \eta} = x_\nu + \frac{a_\nu b_\nu}{a_\nu + b_\nu},$$

$$a_{\nu+1} = a_\nu - \frac{a_\nu b_\nu}{a_\nu + b_\nu} = \frac{a_\nu^2}{a_\nu + b_\nu}.$$

В силу симметрии аналогичное равенство верно для $b_{\nu+1}$, поэтому

$$a_{\nu+1} = \frac{a_\nu^2}{a_\nu + b_\nu}; \quad b_{\nu+1} = \frac{b_\nu^2}{a_\nu + b_\nu}. \qquad \text{(VI. 5)}$$

---

[1]) Мы предполагаем, что $a + b \neq 0$, т. е. $x_0 \neq \frac{\xi + \eta}{2}$, так как в противном случае алгоритм Ньютона не определен. — *Прим. перев.*

Мы покажем теперь, что при $\xi \neq \eta$

$$a_\nu = \frac{a^{2^\nu}(a-b)}{a^{2^\nu}-b^{2^\nu}}, \quad b_\nu = \frac{b^{2^\nu}(a-b)}{a^{2^\nu}-b^{2^\nu}} \quad (\xi \neq \eta), \qquad \text{(VI. 6)}$$

а при $\xi = \eta$, $a = b$

$$a_\nu = b_\nu = \frac{a}{2^\nu} \quad (\xi = \eta). \qquad \text{(VI. 6a)}$$

3. Действительно, при $\xi \neq \eta$ и $\nu = 0$ равенства (VI. 6), очевидно, справедливы. Предположим, что (VI. 6) справедливо при некотором $\nu \geqslant 0$, тогда из (VI. 5) имеем

$$a_{\nu+1} = \frac{a^{2^{\nu+1}}(a-b)^2}{(a^{2^\nu}-b^{2^\nu})^2} : \frac{(a^{2^\nu}+b^{2^\nu})(a-b)}{a^{2^\nu}-b^{2^\nu}} = \frac{a^{2^{\nu+1}}(a-b)}{a^{2^{\nu+1}}-b^{2^{\nu+1}}}.$$

Первое равенство из (VI. 6) доказано по индукции. В силу симметрии второе равенство из (VI. 6) также справедливо.

При $\xi = \eta$, $a_\nu = b_\nu$ формулы (VI. 5) имеют вид

$$a_{\nu+1} = b_{\nu+1} = \frac{a_\nu}{2}.$$

Отсюда сразу следует (VI. 6a).

4. В случае $\xi = \eta$ утверждение нашей теоремы следует непосредственно из (VI. 6a), так как в этом случае

$$a_\nu = \xi - x_\nu \to 0 \quad (\nu \to \infty).$$

Рассмотрим случай $\xi \neq \eta$. Если выполнено условие (VI. 2), т. е. $|a| < |b|$, то из (VI. 6) получаем

$$a_\nu \sim (b-a)\left(\frac{a}{b}\right)^{2^\nu} \to 0 \quad (\nu \to \infty).$$

Если же $|a| = |b|$, но $a \neq b$, то из (VI. 6) следует, что $|a_\nu| = |b_\nu|$ и $a_\nu - b_\nu = a - b$. Поэтому если бы в этом случае алгоритм Ньютона сходился, то и a_ν и b_ν стремились бы к нулю одновременно и мы имели бы $a = b$, вопреки нашему предположению. Теорема доказана.

5. Выясним теперь, насколько выполняются достаточные условия теоремы 7.1 в установленных выше случаях сходимости, т. е. при $a = b$ или $|a| \neq |b|$.

Примем обозначения (VI. 4) и положим $p = \frac{b}{a}$; тогда нашим случаям сходимости соответствуют $p = 1$ или $|p| \neq 1$. Если $|p| \neq 1$, то

$$h_\nu = x_{\nu+1} - x_\nu = \frac{a_\nu b_\nu}{a_\nu + b_\nu} = a^{2^\nu} b^{2^\nu} \frac{a - b}{a^{2^{\nu+1}} - b^{2^{\nu+1}}},$$

а если $p = 1$, то

$$h_\nu = \frac{a}{2^{\nu+1}}.$$

Итак,

$$h_\nu = \begin{cases} \dfrac{a}{2^{\nu+1}} & (p = 1), \\ ap^{2^\nu} \dfrac{p-1}{p^{2^{\nu+1}} - 1} & (|p| \neq 1). \end{cases} \tag{VI. 7}$$

С другой стороны, по формуле (VI. 4) имеем при $|p| \neq 1$

$$f'(x_\nu) = 2x_\nu - \xi - \eta = -(a_\nu + b_\nu) =$$

$$= -(a^{2^\nu} + b^{2^\nu}) \frac{a - b}{a^{2^\nu} - b^{2^\nu}} = -a(p-1)\frac{p^{2^\nu} + 1}{p^{2^\nu} - 1},$$

а при $p = 1$

$$f'(x_\nu) = -\frac{a}{2^{\nu-1}}.$$

Так как при этом число M из теоремы 7.1 равно 2, то получаем при $|p| \neq 1$

$$\frac{2Mh_\nu}{f'(x_\nu)} = -4\frac{p^{2^\nu}(p^{2^\nu} - 1)}{(p^{2^{\nu+1}} - 1)(p^{2^\nu} + 1)} = -4\frac{p^{2^\nu}}{(p^{2^\nu} + 1)^2} =$$

$$= -\frac{4}{(p^{2^{\nu-1}} + p^{-2^{\nu-1}})^2}, \tag{VI. 8}$$

а при $p = 1$

$$\frac{2Mh_\nu}{f'(x_\nu)} = -1. \tag{VI. 8a}$$

6. Итак, в случае $p = 1$ при любом ν имеет место равенство $\left|\frac{2Mh_\nu}{f'(x_\nu)}\right| = 1$, что является предельным случаем достаточного условия теоремы 7.1. Как обстоит дело при $|p| \neq 1$?

В этом случае модуль выражения в левой части формулы (VI.8) стремится к нулю при $\nu \to \infty$. Поэтому, начиная с некоторого ν, условия теоремы 7.1 обязательно будут выполняться. С другой стороны, каким бы большим ни было натуральное число N, мы можем так выбрать число p, что условия теоремы 7.1 не будут выполнены для $\nu = 0, 1, \ldots, N$. Действительно, пусть $p = \varrho e^{i\alpha}$, тогда

$$\left|p^{2^{\nu-1}} + p^{-2^{\nu-1}}\right|^2 = \varrho^{2^\nu} + \varrho^{-2^\nu} + 2\cos 2^\nu \alpha.$$

Возьмем здесь $\alpha = \frac{\pi}{2^N}$, тогда при $\nu = 0, 1, \ldots, N$

$$\left|p^{2^{\nu-1}} + P^{-2^{\nu-1}}\right|^2 \leqslant \varrho^{2^\nu} + \varrho^{-2^\nu} + 2\cos\frac{\pi}{2^N}.$$

Выбрав теперь $\varrho > 1$ достаточно близко к 1, добьемся, чтобы неравенство $\varrho^{2^\nu} + \varrho^{-2^\nu} < 4 - 2\cos(\pi/2^N)$ выполнялось при $\nu = 0, 1, \ldots, N$. Тогда при всех этих ν модуль правой части формулы (VI. 8) будет больше единицы.

7. Выясним, наконец, может ли модуль выражения (VI. 8) быть равным единице при одном или двух последовательных значениях ν. Из тождества $q^2 + \frac{1}{q^2} = \left(q + \frac{1}{q}\right)^2 - 2$ заключаем, что равенства $\left|q^2 + \frac{1}{q^2}\right| = 2$ и $\left|q + \frac{1}{q}\right| = 2$ могут выполняться одновременно лишь при условии, что

$$q^2 + \frac{1}{q^2} = 2; \quad q^2 = 1; \quad q = \pm 1.$$

Так как в (VI. 8) $|p| \neq 1$, то модуль выражения (VI.8) не может равняться единице при двух последовательных значениях ν.

Мы видим, в частности, что если квадратичный многочлен $f(x)$ имеет простые корни, т. е. $p \neq 1$ и если при некотором ν выражение $\left|\frac{2Mh_\nu}{f'(x_\nu)}\right|$ равно 1, то при всех последующих значениях ν это выражение меньше единицы.

ПРИЛОЖЕНИЕ VII

НЕКОТОРЫЕ МОДИФИКАЦИИ И УЛУЧШЕНИЯ МЕТОДА НЬЮТОНА

1. В некоторых руководствах встречается утверждение, что в формуле (6.3) знаменатель $f'(x_\nu)$ можно заменить на $f'(x_1)$, если точка x_1 достаточно близка к корню ζ. Такое утверждение, конечно, ошибочно, так как мы получаем при этом итерационную формулу

$$x_{\nu+1} = x_\nu - cf(x_\nu),$$

дающую при $c \neq \frac{1}{f'(\zeta)}$ только *линейную сходимость*, а не сверхлинейную сходимость, характерную для метода Ньютона.

Приводимая ниже таблица составлена для уравнения $x^3 - 2x - 5 = 0$, которое мы уже исследовали в гл. 3.

$x^3 - 2x - 5 = 0$, $\zeta = 2{,}094551481542326591482386 54$, $x_0 = 2$

I	II
$x_1 = 2{,}1$ $x_2 = 2{,}0945681$ $x_3 = 2{,}09455148172$ — — —	$x_1 = 2{,}1$ $x_2 = 2{,}0939$ $x_3 = 2{,}094627$ $x_4 = 2{,}0945427$ $x_5 = 2{,}0945525$ $x_6 = 2{,}094551363$
III	**IV**
$y_0 = 2{,}1$ $x_1 = 2{,}0939$ $y_1 = 2{,}09455172$ $x_2 = 2{,}09455148136728$	$y_0 = 2{,}1$ $x_1 = 2{,}09456328$ $y_1 = 2{,}0945514816206945426$ $x_2 = 2{,}0945514815423265919$

В столбце I стоят три значения x_1, x_2, x_3, полученные по формуле Ньютона, в столбце II — 6 значений x_ν, полученных по упрощенной формуле, в которой величина $f'(x_\nu)$ заменена на $f'(x_0)$; в обоих случаях $x_0 = 2$.

Сравнивая полученные числа с величиной ζ, видим, что в столбце II погрешность на каждом шаге составляет лишь около 1/10 погрешности предыдущего числа.

2. С другой стороны, может быть все же удобно вычислять $f'(x_\nu)$ не на каждом шаге расчета, *а только на каждом втором шаге*. Это правило можно изложить следующим образом: по ν-му приближению x_ν к ζ находим следующее приближение $x_{\nu+1}$, полагая

$$y_\nu = x_\nu - \frac{f(x_\nu)}{f'(x_\nu)}; \qquad x_{\nu+1} = y_\nu - \frac{f(y_\nu)}{f'(x_\nu)}. \qquad \text{(VII. 1)}$$

Предположим теперь, что последовательность x_ν (а значит и y_ν) сходится[1]) к ζ, и исследуем быстроту сходимости при условии, что $f'(\zeta) \neq 0$.

Из формулы (6.9), заменяя в ней x_0 на x_ν и x_1 на y_ν, получаем

$$\frac{y_\nu - \zeta}{(\zeta - x_\nu)^2} \to \frac{1}{2}\frac{f''(\zeta)}{f'(\zeta)}, \quad (x_\nu \to \zeta). \qquad \text{(VII. 2)}$$

С другой стороны, при $y_\nu \to \zeta$,

$$f(y_\nu) = (y_\nu - \zeta) f'(\zeta) + O[(y_\nu - \zeta)^2];$$

поэтому из формул (VI. 1) и (VI. 2) имеем

$$x_{\nu+1} - \zeta = y_\nu - \zeta - \frac{1}{f'(x_\nu)}(y_\nu - \zeta) f'(\zeta) + O[(y_\nu - \zeta)^2],$$

$$\frac{x_{\nu+1} - \zeta}{y_\nu - \zeta} = \frac{f'(x_\nu) - f'(\zeta)}{f'(x_\nu)} + O(y_\nu - \zeta) =$$

$$= \frac{f''(\xi)}{f'(x_\nu)}(x_\nu - \zeta) + O[(x_\nu - \zeta)^2], \ \xi \in (x_\nu, \zeta).$$

Отсюда, учитывая, что $f'(\zeta) \neq 0$, находим

$$\frac{x_{\nu+1} - \zeta}{(y_\nu - \zeta)(x_\nu - \zeta)} \to \frac{f''(\zeta)}{f'(\zeta)}, \qquad \text{(VII. 3)}$$

---

[1]) Последовательность (VII.1) может расходиться даже в том случае, когда метод Ньютона сходится (например, для нашего уравнения при $x_0 = 1{,}5$). При выполнении условий Фурье (см. стр. 64) последовательность (VII.1) всегда сходится.— *Прим. перев.*

и по формуле (VII. 2) получаем окончательно

$$\frac{x_{\nu+1}-\zeta}{(x_\nu-\zeta)^3} \to \frac{1}{2}\left(\frac{f''(\zeta)}{f'(\zeta)}\right)^2, \quad (x_\nu \to \zeta). \qquad \text{(VII. 4)}$$

3. Заметим теперь, что вычисление по формулам (VII. 1) требует *трех горнеров*. Поэтому «ценой» шести горнеров мы можем перейти от x_ν к $x_{\nu+2}$, причем

$$\frac{x_{\nu+2}-\zeta}{(x_\nu-\zeta)^9} \to \frac{1}{16}\left(\frac{f''(\zeta)}{f'(\zeta)}\right)^8. \qquad \text{(VII. 5)}$$

С другой стороны, если мы применим трижды формулу (6. 3), то «ценой» шести горнеров перейдем от x_ν к $x_{\nu+3}$, причем, как следует из формулы (6.9),

$$\frac{x_{\nu+3}-\zeta}{(x_\nu-\zeta)^8} \to \left(\frac{f''(\zeta)}{2f'(\zeta)}\right)^7. \qquad \text{(VI. 6)}$$

Мы видим, что наше новое правило (VII. 1) действительно является улучшением правила (6.3), хотя необходимость вычисления $f'(x_\nu)$ с бо́льшим количеством десятичных знаков, чем при классическом методе Ньютона, конечно, является недостатком.

В столбце III приведенной выше таблицы даны значения y_0, x_1, y_1, x_2, полученные по правилу (VII. 1) с начальной точкой $x_0 = 2$. Здесь погрешность приближения x_2 имеет тот же порядок, что и погрешность приближения x_3 в столбце I, хотя из формул (VII. 5) и (VII. 6) следует, что приближение x_2 в столбце III должно давать меньшую погрешность. Дело в том, что здесь $\frac{f''(\zeta)}{f'(\zeta)} \approx 1{,}2$, и поэтому множитель $8\,\frac{f''(\zeta)}{f'(\zeta)}$, отличающий (VII. 5) от (VII. 6), компенсирует влияние множителя $(x_\nu-\zeta)$ до тех пор, пока множитель $(x_\nu-\zeta)$ не будет достаточно мал.

4. С другой стороны, можно пытаться понизить число горнеров в формуле Ньютона, заменяя на каждом втором шаге знаменатель $f'(x_\nu)$ подходящей комбинацией $f(x_\nu)$ и $f(x_{\nu-1})$.

В этом направлении может быть использовано следующее правило.

Начиная с x_0, построим последовательность $\{x_\nu\}$ по формулам

$$y_\nu = x_\nu - \frac{f(x_\nu)}{f'(x_\nu)}, \quad x_{\nu+1} = y_\nu - \frac{f(y_\nu)(y_\nu - x_\nu)}{2f(y_\nu) - f(x_\nu)} \quad (\nu = 0, 1, \ldots). \tag{VII. 7}$$

Мы видим, что переход от x_ν к $x_{\nu+1}$ требует трех горнеров. Если x_ν стремится к ζ и $f'(\zeta) \neq 0$, то, как будет показано в п. 5—7,

$$\frac{x_{\nu+1} - \zeta}{(x_\nu - \zeta)^4} \to \frac{1}{24} \frac{f''(\zeta)}{f'(\zeta)^3} [3f''(\zeta)^2 - 2f'(\zeta) f'''(\zeta)]. \tag{VII. 8}$$

Таким образом, «ценой» трех горнеров мы получаем при этом методе улучшение того же порядка, который при методе Ньютона требует двух шагов, т. е. четырех горнеров [1]).

5. Нам остается доказать, что при $x_0 \to \zeta$

$$\frac{x_1 - \zeta}{(x_0 - \zeta)^4} \to \frac{1}{24} \frac{f''(\zeta)}{f'(\zeta)^3} [3f''(\zeta)^2 - 2f'(\zeta) f'''(\zeta)]. \tag{VII. 9}$$

Положим

$$\begin{aligned} y_0 - x_0 &= h = -\frac{f(x_0)}{f'(x_0)}; \\ x_1 - y_0 &= \varrho = \frac{-f(y_0)\, h}{2f(y_0) - f(x_0)}. \end{aligned} \tag{VII. 10}$$

---

[1]) Формулу (VII.8) можно получить как следствие формулы (11.29). Действительно, если мы заменим в формуле (11.14) x_0 на x_ν и x_1 на y_ν, то получим приближение, совпадающее с $x_{\nu+1}$. Поэтому к $x_{\nu+1}$ применима формула (11.29):

$$\frac{x_{\nu+1} - \zeta}{(x_\nu - \zeta)^2 (y_\nu - \zeta)} \to \frac{3f''(\zeta)^2 - 2f'''(\zeta) f'(\zeta)}{12 f'(\zeta)^2}.$$

Здесь, как следует из формул (6.9) и (VII.7), имеет место

$$y_\nu - \zeta \sim (x_\nu - \zeta)^2 \cdot \frac{f''(\zeta)}{2f'(\zeta)},$$

и мы получаем формулу (VII.8).

Заметим также, что формулы (VII.7) проще для вычислений и определяют более эффективный итерационный процесс, чем формула (11.14). — *Прим. перев.*

Заменяя в формуле (6.9) x_1 на y_0, получаем

$$\frac{\zeta - y_0}{(\zeta - x_0)^2} \to -\frac{f''(\zeta)}{2f'(\zeta)}, \quad \zeta - y_0 = O[(\zeta - x_0)^2]; \quad \text{(VII. 11)}$$

поэтому

$$h = y_0 - x_0 = (\zeta - x_0) - (\zeta - y_0) = \zeta - x_0 - O[(\zeta - x_0)^2],$$

$$\zeta - x_0 \sim h, \quad \text{(VII. 12)}$$

$$\zeta - y_0 \sim -\frac{f''(\zeta)}{2f'(\zeta)} h^2. \quad \text{(VII. 13)}$$

Далее,

$$f(x_0) = f(x_0) - f(\zeta) \sim (x_0 - \zeta) f'(\zeta) \sim -hf'(\zeta). \quad \text{(VII. 14)}$$

и

$$f(y_0) = f(y_0) - f(\zeta) \sim (y_0 - \zeta) f'(\zeta) \sim \frac{f''(\zeta)}{2} h^2. \quad \text{(VII. 15)}$$

6. Покажем теперь, что

$$f(x_1) = O(h^4).$$

С этой целью разложим $f(x_1) = f(y_0 + \varrho)$ по степеням ϱ:

$$f(x_1) = [f(y_0) + \varrho f'(y_0)] + \frac{\varrho^2}{2} f''(\tau); \quad \text{(VII. 16)}$$

здесь $\tau \in (y_0, x_1)$ и потому $\tau \to \zeta$ при $y_0, x_1 \to \zeta$. Из (VII. 10), учитывая (VII. 14) и (VII. 15), получаем

$$\varrho = \frac{-f(y_0) h}{2f(y_0) - f(x_0)} \sim -\frac{f''(\zeta)}{2f'(\zeta)} h^2. \quad \text{(VII. 17)}$$

Итак, последнее слагаемое в (VII. 16) имеет порядок h^4:

$$\frac{\varrho^2}{2} f''(\tau) \sim \frac{f''(\zeta)^3}{8f'(\zeta)^2} h^4. \quad \text{(VII. 18)}$$

Остается оценить первое слагаемое. Из (VII. 10) и (VII. 7) получаем

$$f(y_0) + \varrho f'(y_0) =$$

$$= \frac{f(y_0)}{2f(y_0) - f(x_0)} [2f(y_0) - f(x_0) - hf'(y_0)] \sim \frac{f''(\zeta)}{2f'(\zeta)} hZ,$$

где, в силу соотношения $f(x_0) = -h f'(x_0)$,

$$Z = 2f(x_0 + h) - h[f'(x_0 + h) - f'(x_0)].$$

Разложим Z по степеням h до h^3:

$$Z = 2\left[f(x_0) + hf'(x_0) + \frac{h^2}{2}f''(x_0) + \frac{h^3}{6}f'''(\xi)\right] -$$
$$- h\left[hf''(x_0) + \frac{h^2}{2}f'''(\xi_1)\right] = h^3\left[\frac{f'''(\xi)}{3} - \frac{f'''(\xi_1)}{2}\right];$$

здесь ξ, $\xi_1 \in (x_0, y_0)$ и потому $\xi, \xi_1 \to \zeta$ при $x_0, y_0 \to \zeta$. Итак,

$$f(y_0) + \varrho f'(y_0) \sim \frac{f''(\zeta)}{2f'(\zeta)}\, h \cdot \left[-h^3\frac{f'''(\zeta)}{6}\right]. \qquad \text{(VII. 19)}$$

7. Из формул (VII. 16), (VII. 18), (VII. 19) следует, что

$$f(x_1) \sim \left[\frac{f''(\zeta)^3}{8f'(\zeta)^2} - \frac{f''(\zeta)f'''(\zeta)}{12f'(\zeta)}\right]h^4. \qquad \text{(VII.20)}$$

Мы видим, что действительно $f(x_1) = O(h^4)$.

С другой стороны, из $f(x_1) = f(x_1) - f(\zeta) \sim (x_1 - \zeta)f'(\zeta)$ следует, что

$$x_1 - \zeta \sim \frac{f(x_1)}{f'(\zeta)};$$

поэтому из формул (VII. 12) и (VII. 20) окончательно получаем (VII. 9).

8. В столбце IV таблицы на стр. 171 числа y_0, x_1, y_1, x_2 получены по формуле (VII. 7) с начальной точкой $x_0 = 2$. Мы видим, что «ценой» шести горнеров в столбце IV достигнут намного лучший результат, чем полученный «той же ценой» в столбце I.

С другой стороны, не надо забывать, что, применяя формулу (VII. 7), мы должны вести вычисления $f(x_\nu)$ с двойным числом десятичных знаков по сравнению с тем, что нужно для вычисления y_ν. Кроме того, самокорректирующее свойство метода Ньютона не играет здесь такой роли, как в самом методе Ньютона.

ПРИЛОЖЕНИЕ VIII

ОКРУГЛЕНИЕ ПРИ ОБРАТНОЙ ИНТЕРПОЛЯЦИИ

1. Если мы многократно применяем обратную интерполяцию с фиксированным числом n узлов на каждом шаге, то, как это было показано в гл. 13, п. 2, погрешность приближения x_{n+1} есть $O\left(\prod_{\nu=1}^{n} |\zeta - x_\nu|\right)$ в силу формулы (13.5). Очевидно, точность вычисления величин $y_\nu = f(x_\nu)$ должна быть такой, чтобы эта теоретическая погрешность не была ухудшена. Целью данного приложения является определение допустимых погрешностей в вычислении y_ν.

Обозначим выражение, стоящее в правой части формулы (13.1), как функцию от $x_1, \ldots, x_n; y_1, \ldots, y_n$ через $X = X(x_1, \ldots, x_n;\ y_1, \ldots, y_n)$.

Пусть погрешность вычисления y_ν равна δ_ν, тогда результирующая погрешность вычисления x_{n+1} есть

$$\sum_{\nu=1}^{n} \delta_\nu \frac{\partial X}{\partial y_\nu},$$

где производные $\frac{\partial X}{\partial y_\nu}$ должны быть взяты в точке $(x_1, \ldots, x_n;\ \vartheta_1 y_1, \ldots, \vartheta_n y_n)$, $|\vartheta_\nu| \leqslant 1$ $(\nu = 1, \ldots, n)$. Так как $x_\nu \to \zeta$ и $y_\nu \to 0$, то важно знать порядок величин $\frac{\partial X}{\partial y_\nu}$ при $y_\nu \to 0$, $x_\nu \to \zeta$ $(\nu = 1, \ldots, n)$.

2. В этом исследовании мы предполагаем с самого начала, что $\zeta = 0$ и $f'(0) = 1$. Очевидно, такое предположение можно сделать без потери общности, так как мы всегда можем вычесть ζ из x и умножить $f(x)$ на соот-

ветствующую постоянную. При этих предположениях $\frac{y_\nu}{x_\nu} \to 1$. С другой стороны, так как при любом ν величина $x_{\nu+1}$ дает существенно лучшее приближение, чем x_ν, естественно предположить, что $\frac{y_{\nu+1}}{y_\nu} \to 0$. Следующая лемма дает частное решение стоящей перед нами проблемы.

3. Лемма 1. Пусть величины x_ν, y_ν ($\nu = 1, \ldots, n$; $n \geqslant 2$) стремятся к нулю таким образом, что

$$\frac{x_\nu}{y_\nu} \to 1 \quad (\nu = 1, \ldots, n), \quad \frac{y_{\nu+1}}{y_\nu} \to 0 \quad (\nu = 1, \ldots, n-1). \qquad \text{(VIII. 1)}$$

Положим $\Phi(y) \equiv \prod_{\nu=1}^{n} (y - y_\nu)$ и построим

$$X = (-1)^{n-1} y_1 \ldots y_n \sum_{\nu=1}^{n} \frac{x_\nu}{y_\nu} \frac{1}{\Phi'(y_\nu)} \qquad \text{(VIII. 2)}$$

и

$$D_k = \frac{1}{y_1 \ldots y_n} \frac{\partial X}{\partial y_k} \quad (1 \leqslant k \leqslant n). \qquad \text{(VIII. 3)}$$

Тогда

$$D_n \sim \frac{-1}{y_1 \ldots y_n},$$

$$D_{n-1} \sim \frac{1}{y_1 \ldots y_{n-2} y_{n-1}^2}, \qquad \text{(VIII. 4)}$$

$$D_{n-2} \sim \frac{-1}{y_1 \ldots y_{n-3} y_{n-2}^3} \quad (n \geqslant 3), \qquad \text{(VIII. 5)}$$

$$D_k = o\left(\frac{1}{y_1 \ldots y_{n-2} y_k^2}\right) \quad (1 \leqslant k < n-2, \quad n \geqslant 4). \qquad \text{(VIII. 6)}$$

4. Доказательство. При фиксированном k, $1 \leqslant k \leqslant n$, положим

$$\Phi(y) = (y - y_k) G(y), \quad G(y) = \prod_{\substack{\nu=1 \\ \nu \neq k}}^{n} (y - y_\nu). \qquad \text{(VIII. 7)}$$

Тогда из (VIII. 1) следует, что

$$G(y_k) \sim (-1)^{k-1} y_1 \dots y_{k-1} y_k^{n-k}. \qquad \text{(VIII. 8)}$$

Возьмем теперь логарифмическую производную от $G(y)$; тогда при $k < n$ имеем

$$\frac{G'(y_k)}{G(y_k)} = \sum_{\nu=1}^{k-1} \frac{1}{y_k - y_\nu} + \sum_{\nu=k+1}^{n} \frac{1}{y_k - y_\nu} =$$

$$= -\sum_{\nu=1}^{k-1} \frac{1+o(1)}{y_\nu} + \frac{1}{y_k} \sum_{\nu=k+1}^{n} [1+o(1)],$$

$$y_k \frac{G'(y_k)}{G(y_k)} \to n-k.$$

При $k = n$ имеем

$$\frac{G'(y_n)}{G(y_n)} = -\sum_{\nu=1}^{n-1} \frac{1}{y_\nu} \frac{1}{1-\frac{y_n}{y_\nu}} \sim -\frac{1}{y_{n-1}};$$

следовательно,

$$\frac{G'(y_k)}{G(y_k)} \sim \begin{cases} \dfrac{n-k}{y_k} & (k<n), \\ \dfrac{-1}{y_{n-1}} & (k=n). \end{cases}$$

Поэтому, учитывая формулу (VIII. 8), получаем

$$\frac{G'(y_k)}{G(y_k)^2} \sim \begin{cases} (-1)^{k-1} \dfrac{n-k}{y_1 \dots y_{k-1} y_k^{n-k+1}} & (k<n), \\ \dfrac{(-1)^n}{y_1 y_2 \dots y_{n-2} y_{n-1}^2} & (k=n). \end{cases} \qquad \text{(VIII. 9)}$$

5. С другой стороны, при $\nu \neq k$ имеем

$$G'(y_\nu) = \prod_{\substack{\sigma=1 \\ \sigma \neq k, \nu}}^{n} (y_\nu - y_\sigma);$$

поэтому

$$G'(y_\mu) \sim (-1)^{\mu-1} y_1 \dots y_{\mu-1} y_\mu^{n-\mu-1} \quad (\mu < k), \qquad \text{(VIII. 10)}$$

$$G'(y_\lambda) \sim (-1)^{\lambda} \frac{y_1 \dots y_{\lambda-1} y_\lambda^{n-\lambda}}{y_k} \quad (\lambda > k). \qquad \text{(VIII. 11)}$$

Далее, по формуле (VIII. 2) имеем

$$X = \sum_{\nu=1}^{n} U_\nu, \quad U_\nu = (-1)^{n-1} y_1 \dots y_n \frac{x_\nu}{y_\nu} \frac{1}{\Phi'(y_\nu)}.$$

При этом из формул (VIII. 7) следует, что

$$U_k = (-1)^{n-1} y_1 \dots y_n \frac{x_k}{y_k} \frac{1}{G(y_k)},$$

$$U_\nu = (-1)^{n-1} y_1 \dots y_n \frac{x_\nu}{y_\nu} \frac{1}{G'(y_\nu)(y_\nu - y_k)} \quad (\nu \neq k).$$

Дифференцируя по y_k, получаем

$$\frac{\partial U_k}{\partial y_k} = y_1 \dots y_n \frac{x_k}{y_k} \frac{(-1)^n G'(y_k)}{G(y_k)^2},$$

а при $\nu \neq k$

$$\frac{\partial U_\nu}{\partial y_k} = \frac{y_1 \dots y_n}{y_k} \frac{x_\nu}{y_\nu} \frac{(-1)^{n-1}}{G'(y_\nu)} \frac{\partial}{\partial y_k}\left(\frac{y_k}{y_\nu - y_k}\right) =$$

$$= \frac{y_1 \dots y_n}{y_k} \frac{x_\nu}{y_\nu} \frac{(-1)^{n-1}}{G'(y_\nu)} \frac{y_\nu}{(y_\nu - y_k)^2};$$

поэтому

$$D_k = \frac{1}{y_1 \dots y_n} \frac{\partial X}{\partial y_k} = \sum_{\nu=1}^{n} T_\nu, \qquad \text{(VIII. 12)}$$

$$T_k = \frac{(-1)^n x_k G'(y_k)}{y_k G(y_k)^2}; \quad T_\nu = \frac{(-1)^{n-1} x_\nu}{y_k G'(y_\nu)(y_\nu - y_k)^2} \quad (\nu \neq k).$$

6. Из (VIII. 9) следует, что

$$\left.\begin{aligned} T_k &\sim \frac{(-1)^{n-k+1}(n-k)}{y_1 \dots y_{k-1} y_k^{n-k+1}} \quad (k < n), \\ T_k &\sim -\frac{1}{y_1 \dots y_{n-2} y_{n-1}^2} \quad (k = n), \end{aligned}\right\} \qquad \text{(VIII. 13)}$$

а из формул (VIII. 10) и (VIII. 11) при $\mu < k$ и $\lambda > k$

$$T_\mu \sim \frac{(-1)^{n-1} y_\mu}{y_k (-1)^{\mu-1} y_1 \dots y_{\mu-1} y_\mu^{n-\mu-1} y_\mu^2} =$$

$$= \frac{(-1)^{n-\mu}}{y_1 \dots y_{\mu-1} y_\mu^{n-\mu} y_k} \quad (\mu < k), \qquad \text{(VIII. 14)}$$

$$T_\lambda \sim \frac{(-1)^{n-1} y_\lambda y_k}{y_k (-1)^\lambda y_1 \dots y_{\lambda-1} y_\lambda^{n-\lambda} y_k^2} =$$

$$= \frac{(-1)^{n-\lambda-1}}{y_1 \dots y_{\lambda-1} y_\lambda^{n-\lambda-1} y_k^2} \quad (\lambda > k). \qquad \text{(VIII. 15)}$$

При $k < n$ из формул (VIII. 13) и (VIII. 14) получаем

$$\frac{T_\mu}{T_k} \sim \frac{(-1)^{k-\mu-1}}{n-k} \left(\frac{y_\mu}{y_\mu}\right) \left(\frac{y_{\mu+1}}{y_\mu}\right) \dots \left(\frac{y_{k-1}}{y_\mu}\right) \left(\frac{y_k}{y_\mu}\right)^{n-k},$$

поэтому

$$\frac{T_\mu}{T_k} \to 0 \quad (\mu < k < n). \qquad \text{(VIII. 16)}$$

Далее, из (VIII. 13), (VIII. 14), (VIII. 15) при $k = n$, $k = n - 1$ или $k \leqslant n - 2$ имеем

$$\frac{T_n}{T_{n-1}} \sim -\frac{y_n}{y_{n-1}} \to 0 \qquad \text{(VIII. 17)}$$

7. Из формулы (VIII. 14) при $k = n$ и $\mu < n - 1$ имеем

$$\frac{T_\mu}{T_{n-1}} \sim (-1)^{n-\mu-1} \left(\frac{y_\mu}{y_\mu}\right) \left(\frac{y_{\mu+1}}{y_\mu}\right) \dots \left(\frac{y_{n-1}}{y_\mu}\right) \to 0$$

$$(\mu < n - 1 < k).$$

Из этого соотношения и (VIII. 17) заключаем, что при $k = n$ в (VIII. 12) доминирует член T_{n-1}, поэтому в силу эквивалентности (VIII. 14)

$$D_n \sim T_{n-1} \sim \frac{-1}{y_1 \dots y_n}.$$

Первое из соотношений (VIII. 4) доказано.

8. При $k = n - 1$ из (VIII. 16) и (VIII. 17) заключаем, что снова доминирует член T_{n-1} и в силу (VIII. 13)

$$D_{n-1} \sim T_{n-1} \sim \frac{1}{y_1 \dots y_{n-2} y_{n-1}^2}.$$

Второе из соотношений (VIII. 4) доказано.

9. При $k = n-2$ из формулы (VIII. 15) с $\lambda = n - 1$ и из формулы (VIII. 13) с $k = n - 2$ получаем

$$\frac{T_{n-2}}{T_{n-1}} \sim \frac{(-1)^3 2 y_1 \dots y_{n-2} y_{n-2}^2}{y_1 \dots y_{n-3} y_{n-2}^3} = -2, \quad \frac{T_{n-2}}{T_{n-1}} \to -2.$$

Отсюда и из формул (VIII. 16) и (VIII. 17) следует, что

$$D_{n-2} \sim -T_{n-1} \sim \frac{-1}{y_1 \dots y_{n-3} y_{n-2}^3}.$$

Формула (VIII. 5) доказана.

10. Рассмотрим теперь случай $k < n-2$. При этом из (VIII. 13) и (VIII. 15) с $\lambda = n-1$ получаем

$$\frac{T_k}{T_{n-1}} \sim (-1)^{n-k+1}(n-k)\left(\frac{y_k}{y_k}\right)\left(\frac{y_{k+1}}{y_k}\right)\dots\left(\frac{y_{n-2}}{y_k}\right) \to 0. \tag{VIII. 18}$$

Для $\lambda < n-1$ получаем из (VIII. 15)

$$\frac{T_\lambda}{T_{n-1}} \sim \frac{(-1)^{n-\lambda+1} y_1 \dots y_{n-2} y_k^2}{y_1 \dots y_{\lambda-1} y_\lambda^{n-\lambda-1} y_k^2} =$$

$$= (-1)^{n-\lambda+1}\left(\frac{y_\lambda}{y_\lambda}\right)\left(\frac{y_{\lambda+1}}{y_\lambda}\right)\dots\left(\frac{y_{n-2}}{y_\lambda}\right).$$

Это выражение стремится к нулю при $\lambda < n-2$ и равно -1 при $\lambda = n-2$. Поэтому из (VIII. 16), (VIII. 17) и (VIII. 18) следует, что

$$T_\nu = o(T_{n-1}) \quad (\nu < n-2,\ \nu = n)$$

и

$$T_{n-2} + T_{n-1} = o(T_{n-1}).$$

Наконец, в силу формулы (VIII. 12)

$$D_k = o(T_{n-1}) = o\left(\frac{1}{y_1 \dots y_{n-2} y_k^2}\right) \quad (k < n-2).$$

Лемма 1 доказана.

11. Используя формулы (VIII. 3)—(VIII. 6), приходим к выводу, что для выполнения соотношения

$$\sum_{\nu=1}^{n} \delta_\nu \frac{\partial X}{\partial y_\nu} = O(y_1 \dots y_n)$$

достаточно, чтобы погрешности δ_ν были подчинены следующим условиям:

$$\left.\begin{aligned}\delta_n &= O(y_1 \dots y_n),\\ \delta_{n-1} &= O\left(y_1 \dots y_{n-2}\, y_{n-1}^2\right),\\ \delta_{n-2} &= O\left(y_1 \dots y_{n-3}\, y_{n-2}^3\right),\end{aligned}\right\} \qquad \text{(a)}$$

$$\delta_k = O\left(y_1 \dots y_{n-2}\, y_k^2\right) \quad (k = 1, \dots, \quad n-3). \qquad \text{(b)}$$

Результаты, содержащиеся в (a), вполне удовлетворительны, так как величины y_ν, входящие в эти оценки, можно считать известными к моменту определения соответствующей погрешности δ_ν.

Напротив, оценки (b) не дают «истинного» порядка величин δ_k и требуют знания y_ν с $\nu > k$.

Однако, принимая дополнительное условие, что $f^{(n-2)}(x)$ существует и непрерывна в J_x, мы можем получить и для δ_k с $k < n-2$ оценки, дающие «наилучший» порядок величин и зависящие только от $y_1, \dots, y_k$. Конечно, потребность в таких оценках возникает только при $n \geqslant 4$. Так как при нашем новом условии $\varphi^{(n-2)}(w)$ существует и непрерывна в окрестности начала координат, то, полагая $\varphi'(0) = 1$, получаем

$$\varphi(w) = w + \sum_{\nu=2}^{n-3} a_\nu w^\nu + O(w^{n-2}),$$

откуда при $w = y_\lambda$

$$x_\lambda = y_\lambda + \sum_{\nu=2}^{n-3} a_\nu y_\lambda^\nu + O\left(y_\lambda^{n-2}\right) \quad (1 < \lambda \leqslant n-3).$$

Следующая лемма дает решение стоящей перед нами проблемы.

12. Лемма 2. Пусть, при условиях леммы 1, $n \geqslant 4$, $1 \leqslant k < n-2$, и для любого λ с $k < \lambda \leqslant n$

$$x_\lambda = y_\lambda + \sum_{\nu=2}^{n-k-2} a_\nu y_\lambda^\nu + O\left(y_\lambda^{n-k-1}\right) \quad (k < \lambda \leqslant n), \quad \text{(VIII. 19)}$$

где a_ν — постоянные, не зависящие от λ. Тогда вместо (VIII. 6) имеем оценку

$$D_k \sim \frac{(-1)^{n-k+1}}{y_1 \cdots y_{k-1}\, y_k^{n-k+1}} \quad (1 \leqslant k \leqslant n-3). \quad \text{(VIII. 20)}$$

13. Доказательство. Установим прежде всего некоторые соотношения, вытекающие из условий леммы 1. Положим $\eta = y_k$; тогда из формулы (VIII. 11) получаем при $\lambda > k$ и любом целом ν

$$\frac{y_\lambda^\nu}{\eta(\eta - y_\lambda)^2 G'(y_\lambda)} \sim \frac{(-1)^\lambda y_\lambda^{\nu-n+\lambda}}{\eta^2 y_1 \ldots y_k y_{k+1} \ldots y_{\lambda-1}}.$$

Порядок правой части не уменьшится, если мы заменим множители $y_{k+1}, \ldots, y_{\lambda-1}$ на y_λ, поэтому

$$\frac{y_\lambda^\nu}{\eta(\eta - y_\lambda)^2 G'(y_\lambda)} = O\left(\frac{y_\lambda^{\nu+1-(n-k)}}{y_1 \ldots y_k \eta^2}\right) = O\left(\frac{\eta^{\nu-1-(n-k)}}{y_1 \ldots y_k}\right).$$

Мы предполагаем здесь, что $\nu + 1 \geqslant n - k$. Воспользовавшись первой из эквивалентностей (VIII. 13), получаем

$$\frac{y_\lambda^\nu}{\eta(\eta - y_\lambda)^2 G'(y_\lambda)} = O(\eta^{\nu-1} T_k) \qquad (\nu + 1 \geqslant n - k,\ k < \lambda \leqslant n). \tag{VIII. 21}$$

14. Положим $G(y) = G_1(y) G_2(y)$, где

$$G_1(y) = \prod_{\mu=1}^{k-1} (y - y_\mu); \quad G_2(y) = \prod_{\lambda=k+1}^{n} (y - y_\lambda).$$

Очевидно, что

$$G'(y_\lambda) = G_1(y_\lambda) G_2'(y_\lambda) \qquad (k < \lambda \leqslant n)$$

и

$$G_1(y_\lambda) \sim (-1)^{k-1} y_1 \ldots y_{k-1}, \quad G_2(\eta) \sim \eta^{n-k},$$
$$\eta \frac{G_2'(\eta)}{G_2(\eta)} = \sum_{\lambda=k+1}^{n} \frac{\eta}{\eta - y_\lambda} \to n - k. \tag{VIII. 22}$$

Поэтому из формулы (VIII. 21) получаем

$$\frac{y_\lambda^\nu}{(\eta - y_\lambda)^2 G_2'(y_\lambda)} = O(\eta G_1(y_\lambda) \eta^{\nu-1} T_k) = O(y_1 \ldots y_k \eta^{\nu-1} T_k)$$
$$(\nu + 1 \geqslant n - k,\ k < \lambda \leqslant n). \tag{VIII. 23}$$

15. Рассмотрим выражение

$$K_\nu \equiv \sum_{\lambda=k+1}^{n} \frac{y_\lambda^\nu}{(\eta - y_\lambda)^2 G_2'(y_\lambda)} =$$

$$= -\frac{\partial}{\partial \eta} \frac{1}{G_2(\eta)} \sum_{\lambda=k+1}^{n} \frac{y_\lambda^\nu G_2(\eta)}{(\eta - y_\lambda) G_2'(y_\lambda)}. \qquad \text{(VIII. 24)}$$

При $1 \leqslant \nu < n-k$ сумма, стоящая в правой части, равна η^ν в силу интерполяционной формулы Лагранжа, поэтому из формулы (VIII. 22) следует, что

$$K_\nu = -\frac{\partial}{\partial \eta} \frac{\eta^\nu}{G_2(\eta)} = \frac{\eta^{\nu-1}}{G_2(\eta)} \left[\eta \frac{G_2'(\eta)}{G_2(\eta)} - \nu \right] \sim \frac{n-k-\nu}{\eta^{(n-k+1)-\nu}},$$

и, в частности,

$$K_1 \sim (n-k-1)\,\eta^{k-n}, \quad K_\nu = O(\eta^{\nu-(n-k)-1}) \quad (\nu > 1).$$

Разделив эти соотношения на T_k и воспользовавшись эквивалентностью (VIII. 13), получим окончательно

$$\frac{K_1}{T_k} \sim (-1)^{n-k-1} \frac{n-k-1}{n-k} y_1 \ldots y_k, \qquad \text{(VIII. 25)}$$

$$\frac{K_\nu}{T_k} = O(y_1 \ldots y_k \eta^{\nu-1}) \quad (\nu > 1). \qquad \text{(VIII. 26)}$$

Последняя оценка доказана нами для $\nu < n-k$. Однако она остается справедливой и при $\nu \geqslant n-k$, как это следует из формул (VIII. 23) и (VIII. 24).

16. Введем теперь величины

$$S_\nu \equiv (-1)^{n-1} \sum_{\lambda=k+1}^{n} \frac{y_\lambda^\nu}{\eta(\eta - y_\lambda)^2 G'(y_\lambda)} \qquad \text{(VIII. 27)}$$

$$(\nu = 1, \ldots, n-k-2).$$

Из формул (VIII. 12) и (VIII. 16) имеем

$$D_k = T_k + \sum_{\lambda=k+1}^{n} T_\lambda + o(T_k).$$

Подставляя в выражение для T_λ величины x_λ, определенные по формуле (VIII. 19), получим

$$D_k = T_k + S_1 + \sum_{\nu=2}^{n-k-2} a_\nu S_\nu + \\ + \sum_{\lambda=k+1}^{n} O\left(\frac{y_\lambda^{n-k-1}}{\eta G'(y_\lambda)(\eta - y_\lambda)^2}\right) + o(T_k).$$

В случае $k = 1$, $n = 4$ первую сумму правой части следует считать равной нулю. Каждый член второй суммы правой части есть $o(T_k)$, как следует из формулы (VIII. 21), примененной к $\nu = n - k - 1$ $(\nu > 1$, так как $k < n - 2)$. Поэтому

$$D_k = T_k + S_1 + \sum_{\nu=2}^{n-k-2} a_\nu S_\nu + o(T_k). \qquad \text{(VIII. 28)}$$

17. В случае $k = 1$ имеем $G_2(y) = G(y)$, поэтому из (VIII. 27) и (VIII. 24) следует, что

$$\frac{S_\nu}{T_k} = (-1)^{n-1} \frac{K_\nu}{\eta T_k}.$$

Поэтому при $k = 1$ получаем из (VIII. 25) и (VIII. 26) соотношения

$$S_\nu = o(T_k) \quad (\nu > 1), \qquad \text{(VIII. 29)}$$

$$S_1 \sim -\left(1 - \frac{1}{n-k}\right) T_k. \qquad \text{(VIII. 30)}$$

Мы докажем теперь, что эти соотношения справедливы также при $2 \leqslant k < n - 2$.

18. Из (VIII. 27) следует, что

$$(-1)^{n-k} y_1 \ldots y_k S_\nu = \\ = \sum_{\lambda=k+1}^{n} \prod_{\mu=1}^{k-1} \frac{1}{1 - y_\lambda/y_\mu} \cdot \frac{y_\lambda^\nu}{G_2'(y_\lambda)(\eta - y_\lambda)^2}. \qquad \text{(VIII. 31)}$$

С другой стороны, при любом $\lambda > k$ и любом $\mu < k$ имеем

$$\frac{1}{1 - y_\lambda/y_\mu} = \sum_{\sigma=0}^{n} y_\mu^{-\sigma} y_\lambda^\sigma + O\left(y_\lambda^{n+1} y_\mu^{-n-1}\right),$$

где коэффициент при каждой степени $y_\lambda^\sigma\,(\sigma > 0)$ есть $y_\mu^{-\sigma} = o(\eta^{-\sigma})$. Поэтому

$$\prod_{\mu=1}^{k-1} \frac{1}{1 - y_\lambda/y_\mu} = 1 + \sum_{\sigma=1}^{n+1} F_\sigma\, y_\lambda^\sigma, \qquad \text{(VIII. 32)}$$

где

$$F_\sigma = o(\eta^{-\sigma}) \qquad (\sigma = 1, 2, \ldots, n+1). \qquad \text{(VIII. 33)}$$

Подставляя (VIII. 32) в (VIII. 31), получаем на основании формулы (VIII. 24)

$$(-1)^{n-k} y_1 \ldots y_k S_\nu = K_\nu + \sum_{\sigma=1}^{n+1} F_\sigma K_{\nu+\sigma}. \qquad \text{(VIII. 34)}$$

Вследствие (VIII. 26) каждый член суммы в правой части есть

$$o(\eta^{-\sigma} y_1 \ldots y_k \eta^{\nu+\sigma-1} T_k) = o(y_1 \ldots y_k \eta^{\nu-1} T_k).$$

19. Из (VIII. 34) и (VIII. 26) следует теперь, что

$$S_\nu = o(T_k) \qquad (\nu > 1),$$

$$S_1 = (-1)^{n-k} \frac{K_1}{y_1 \ldots y_k} + o(T_k).$$

Воспользовавшись эквивалентностью (VIII. 25), заключаем, что формулы (VIII. 29) и (VIII. 30) справедливы и в общем случае. Теперь формула (VIII. 20) непосредственно следует из (VIII. 28), (VIII. 30) и (VIII. 13). Лемма 2 доказана.

20. На основании леммы 2 получаем теперь для всех δ_k $(1 \leqslant k \leqslant n)$ условия:

$$\delta_k = O(y_1 \ldots y_{k-1} y_k^{n-k+1})\,(1 \leqslant k \leqslant n), \qquad \text{(c)}$$

которые можно считать вполне удовлетворительными в том же смысле, что и условия (a).

ПРИЛОЖЕНИЕ IX

УЛУЧШЕНИЕ ИТЕРАЦИЙ СО СВЕРХЛИНЕЙНОЙ СХОДИМОСТЬЮ

1. Мы рассмотрим здесь некоторую последовательность z_ν ($\nu = 1, 2, \ldots$), сходящуюся к ζ. Предположим, что при некотором $t > 1$ существует конечный и отличный от нуля предел:

$$\lim_{\nu\to\infty} \frac{|z_{\nu+1} - \zeta|}{|z_\nu - \zeta|^t} = \alpha \quad (\alpha \neq 0,\ \alpha \neq \infty). \qquad \text{(IX. 1)}$$

Если мы хотим остановиться на z_{n+1}, то, как будет показано ниже, степень приближения к ζ можно улучшить, заменив z_{n+1} на

$$Z = z_{n+1} - \frac{|z_n - z_{n+1}|^{t+1}}{|z_{n-1} - z_n|^t} \operatorname{sign}(z_{n+1} - \zeta)^{1)}. \qquad \text{(IX. 2)}$$

В случае, когда мы имеем более точные данные о быстроте сходимости в (IX. 1), улучшение, получаемое по формуле (IX. 2), может быть соответственно уточнено.

Формула (IX.2) дает улучшение и тогда, когда последовательность z_ν задана *итерационной формулой первого порядка:*

$$z_{\nu+1} = \varphi(z_\nu),$$

и в более общем случае, когда $z_{\nu+1}$ получено *итерацией порядка* k:

$$z_{\nu+1} = \varphi(z_\nu, z_{\nu-1}, \ldots, z_{\nu-k+1}).$$

Наши результаты справедливы независимо от того, каким способом получена последовательность z_ν.

---

[1]) Символ $\operatorname{sign} x$ означает $\frac{x}{|x|}$. Конечно, этот символ определен только при $x \neq 0$.

2. Очевидно, что условие (IX. 1) можно записать в виде

$$|z_{n+1}-\zeta| = \alpha |z_n-\zeta|^t (1+\varepsilon_n),$$
$$|z_n-\zeta| = \alpha |z_{n-1}-\zeta|^t (1+\varepsilon_{n-1}), \qquad \text{(IX. 3)}$$

где $\varepsilon_n \to 0$, $\varepsilon_{n-1} \to 0$ при $n \to \infty$. Положим

$$\Delta_n = \max(|\varepsilon_n|, |\varepsilon_{n-1}|, |z_{n-1}-\zeta|^{t-1}). \qquad \text{(IX. 4)}$$

Мы докажем, что в то время, как приближение, даваемое z_{n+1}, характеризуется соотношением

$$\frac{|z_{n+1}-\zeta|}{\alpha^{t+1}|z_{n-1}-\zeta|^{t^2}} \to 1 \qquad (n \to \infty), \qquad \text{(IX. 5)}$$

для приближения Z справедлива оценка

$$\frac{|Z-\zeta|}{\alpha^{t+1}|z_{n-1}-\zeta|^{t^2}} = O(\Delta_n) \quad (n \to \infty). \qquad \text{(IX. 6)}$$

3. Положим

$$|z_{n-1}-\zeta| = \delta.$$

Из (IX. 3) и (IX. 4) следует, что при $n \to \infty$

$$\frac{|z_{n-1}-z_n|}{\delta} = 1 + O\left(\frac{z_n-\zeta}{\delta}\right) = 1 + O(\delta^{t-1}) = 1 + O(\Delta_n),$$
$$|z_{n-1}-z_n|^t = \delta^t [1 + O(\Delta_n)]. \qquad \text{(IX. 7)}$$

Далее, из (IX. 3) и (IX. 4) следует также, что

$$|z_n - z_{n+1}| = |z_n-\zeta| \left|1 - \frac{z_{n+1}-\zeta}{z_n-\zeta}\right| =$$
$$= \alpha\delta^t [1 + O(|z_n-\zeta|^{t-1})](1+\varepsilon_{n-1}) = \qquad \text{(IX. 8)}$$
$$= \alpha\delta^t [1 + O(\delta^{t-1})][1 + O(\Delta_n)],$$
$$|z_n - z_{n+1}| = \alpha\delta^t [1 + O(\Delta_n)].$$

Из (IX. 7) и (IX. 8) имеем

$$\frac{|z_{n+1}-z_n|^{t+1}}{|z_n-z_{n-1}|^t} = \alpha^{t+1}\delta^{t^2}[1 + O(\Delta_n)]. \qquad \text{(IX. 9)}$$

4. С другой стороны, дважды применяя (IX. 3), получаем

$$|z_{n+1}-\zeta| = \alpha |z_n-\zeta|^t [1 + O(\Delta_n)] =$$
$$= \alpha^{t+1}\delta^{t^2}[1 + O(\Delta_n)]. \qquad \text{(IX. 10)}$$

Отсюда непосредственно следует (IX. 5). Положим

$$s_{n+1} = \operatorname{sign}(z_{n+1} - \zeta);$$

тогда из (IX. 9) и (IX. 10) имеем

$$Z - z_{n+1} = -\frac{|z_n - z_{n+1}|^{t+1}}{|z_{n-1} - z_n|^t} s_{n+1} =$$
$$= -s_{n+1}\alpha^{t+1}\delta^{t^2}[1 + O(\Delta_n)],$$
$$z_{n+1} - \zeta = s_{n+1}\alpha^{t+1}\delta^{t^2}[1 + O(\Delta_n)].$$

Разделив обе эти формулы на $\alpha^{t+1}\delta^{t^2}$ и сложив, получаем:

$$\frac{Z - \zeta}{\alpha^{t+1}\delta^{t^2}} = O(\Delta_n),$$

что совпадает с (IX. 6).

5. Если, в частности, выполнено соотношение

$$\varepsilon_n,\ \varepsilon_{n-1} = O(\delta^p), \quad p > 0, \tag{IX. 11}$$

то, полагая

$$\min(p, t-1) = d, \tag{IX. 12}$$

имеем, очевидно, $\Delta_n = O(\delta^d)$; поэтому (IX. 6) можно заменить на

$$Z - \zeta = O(\delta^{t^2+d}). \tag{IX. 13}$$

Обычно $d=1$, так что применение (IX. 2) дает улучшение на 25% при $t=2$, на 11,1% при $t=3$.

6. Для того чтобы применять аппроксимацию (IX. 2), необходимо знать величину $\operatorname{sign}(z_{n+1} - \zeta)$. Эта величина бывает известна во многих случаях, в частности, когда все z_ν вещественны и при некотором целом t существует конечный и отличный от нуля предел отношения

$$\frac{z_{\nu+1} - \zeta}{(z_\nu - \zeta)^t}.$$

Рассмотрим, например, метод Ньютона, при котором, как следует из соотношения (6.9),

$$\frac{x_{\nu+1} - \zeta}{(x_\nu - \zeta)^2} \to \frac{1}{2}\frac{f''(\zeta)}{f'(\zeta)}.$$

Если при этом все x_ν — вещественные числа, то можно применять аппроксимацию (IX. 2), причем не только в конце

расчета, но также и *после каждого шага*, проведенного методом Ньютона. В этом случае имеют место соотношения (IX. 12) и (IX. 13) с $d = 1$.

Аналогичное замечание можно сделать о методе Шредера, исследованном в гл. 8, а также о методах, рассмотренных в приложении VII, для которых надо воспользоваться формулами (VII. 4) и (VII. 8).

В случае метода секущих можно воспользоваться формулой (3.19), если все x_ν вещественны; знаки разностей $(x_\nu - \zeta)$ легко определить последовательно из формулы (3.11). Здесь, однако, улучшение будет не очень значительным, так как в силу формулы (12.31) величины ε_ν и $\varepsilon_{\nu-1}$ имеют порядок

$$O\left(t_2^\nu\right) = O\left(\frac{1}{\log \delta_\nu}\right).$$

То же самое можно сказать об итерации, рассмотренной в гл. 5.

Вообще в случае итерации первого порядка $z_{\nu+1} = \varphi(z_\nu)$ определение величины $\operatorname{sign}(z_\nu - \zeta)$ обычно не представляет затруднений; в этом случае можно даже применять аппроксимацию (IX. 2) на каждом шаге итерации.

Во многих случаях это верно также и для итераций конечного порядка k:

$$z_{\nu+1} = \varphi(z_\nu, z_{\nu-1}, \ldots, z_{\nu-k+1}).$$

Но даже и в наиболее общем случае произвольной последовательности z_ν применение (IX. 2) на последнем шаге расчета может дать заметное улучшение.

В предыдущих рассуждениях мы были вынуждены использовать три последовательных приближения z_{n-1}, z_n, z_{n+1} для того, чтобы исключить α, так как α не предполагалось известным. В случае, когда α известно из теоретического исследования, формула (IX. 2) может быть уточнена. Предположим, что при $z_\nu \to \zeta$

$$\frac{|z_{\nu+1} - \zeta|}{|z_\nu - \zeta|^t} = \alpha + O(|z_\nu - \zeta|), \quad t \geqslant 2, \qquad \text{(IX. 14)}$$

где $\alpha \neq 0$, $\alpha \neq \infty$. Так как

$$\frac{z_\nu - \zeta}{z_\nu - z_{\nu+1}} = 1 + O(|z_\nu - \zeta|),$$

то мы получаем из (IX. 14)

$$|z_{\nu+1}-\zeta| = \alpha\,|z_\nu - z_{\nu+1}|^t + O(|z_\nu-\zeta|^{t+1}),$$

$$z_{\nu+1}-\zeta = \alpha\,|z_\nu - z_{\nu+1}|^t \operatorname{sign}(z_{\nu+1}-\zeta) + O(|z_\nu-\zeta|^{t+1}),$$

$$\zeta = z_{\nu+1} - \alpha\,|z_\nu - z_{\nu+1}|^t \operatorname{sign}(z_{\nu+1}-\zeta) + O(|z_\nu-\zeta|^{t+1}).$$

Мы видим, что выражение

$$Z^* = z_{\nu+1} - \alpha\,|z_\nu - z_{\nu+1}|^t \operatorname{sign}(z_{\nu+1}-\zeta) \qquad \text{(IX.15)}$$

дает лучшее приближение к ζ, чем $z_{\nu+1}$[1]).

---

[1]) В случае, когда $1<t<2$, формула (IX.15) также дает улучшение аппроксимации, но при этом

$$Z^*-\zeta = O(|z_\nu-\zeta|^{2t-1}).$$

— *Прим. перев.*

ПРИЛОЖЕНИЕ X

НАХОЖДЕНИЕ КОРНЕЙ УРАВНЕНИЯ $f(z)=0$ ПО КОЭФФИЦИЕНТАМ РАЗЛОЖЕНИЯ $\frac{1}{f(z)}$

1. В этом приложении мы исследуем уравнение

$$f(z) \equiv a_0 + a_1 z + a_2 z^2 + \ldots = 0 \quad (a_0 = 1); \qquad \text{(X. 1)}$$

степенной ряд (X. 1) будем предполагать сходящимся в круге радиуса r, $0 < r \leqslant \infty$; $f(z)$ может оказаться многочленом степени n, если все a_ν с индексами $\nu > n$ равны 0.

2. В дальнейшем мы пользуемся разложением

$$\Phi(z) \equiv \frac{1}{f(z)} = \sum_{\nu=0}^{\infty} P_\nu z^\nu, \qquad \text{(X. 2)}$$

где степенной ряд (X. 2) имеет радиус сходимости $\varrho_0 > 0$. Коэффициенты P_ν разложения (X. 2) могут быть определены последовательно из тождества

$$1 \equiv (a_0 + a_1 z + a_2 z^2 + \ldots)(P_0 + P_1 z + P_2 z^2 + \ldots),$$

полученного перемножением разложений (X. 1) и (X. 2). Из этого тождества находим

$$\begin{aligned} &a_0 P_0 = 1, \\ &a_1 P_0 + a_0 P_1 = 0, \\ &\ldots\ldots\ldots \\ &a_\nu P_0 + a_{\nu-1} P_1 + \ldots + a_0 P_\nu = 0, \\ &\ldots\ldots\ldots\ldots\ldots \end{aligned} \qquad \text{(X. 3)}$$

откуда легко определить все коэффициенты P_ν.

Если, в частности, $f(z)$ — многочлен степени n, то формулы (X. 3) показывают, что при $\nu \geqslant n$ коэффициенты P_ν удовлетворяют линейному разностному уравнению (12.2) с характеристическим многочленом (12.3).

3. Для решения системы (X. 3) с помощью определителей введем

$$D_{1,\nu}=\begin{vmatrix} a_1 & a_0 & \dots & a_{2-\nu} \\ a_2 & a_1 & \dots & a_{3-\nu} \\ \vdots & \vdots & \vdots & \vdots \\ a_\nu & a_{\nu-1} & & a_1 \end{vmatrix} \quad (\nu=1,\ 2,\ \dots), \quad D_{1,0}=1, \qquad (X.\ 4)$$

полагая

$$a_{-\mu}=0 \quad (\mu=1,\ 2,\ \dots). \qquad (X.\ 5)$$

Так как определитель системы первых $\nu+1$ уравнений (X. 3) равен 1, то, определяя из этих уравнений P_ν, получаем формулу, принадлежащую Вронскому (1811):

$$P_\nu=(-1)^\nu D_{1,\nu}. \qquad (X.\ 6)$$

4. Предположим теперь, что уравнение (X. 1) имеет *простой* корень ξ_1, модуль которого меньше r и меньше модулей всех остальных корней этого уравнения. В этом случае, выбрав соответствующую постоянную $\alpha_1 \neq 0$, можно написать

$$\Phi(z)=\frac{\alpha_1\xi_1}{\xi_1-z}+\sum_{\nu=0}^{\infty} b_\nu z^\nu. \qquad (X.\ 7)$$

где степенной ряд $\sum b_\nu z^\nu$ имеет радиус сходимости $\varrho_1>\varrho_0$. При любом положительном q, удовлетворяющем условию

$$|\xi_1|<\frac{1}{q}<\varrho_1, \qquad (X.\ 8)$$

имеем $b_\nu=o(q^\nu)$, $|\xi_1|q<1$; поэтому

$$P_\nu=\alpha_1\xi_1^{-\nu}+o(q^\nu) \qquad (X.\ 9)$$

и, следовательно,

$$\frac{P_{\nu-1}}{P_\nu}=\xi_1+o(|\xi_1|^\nu q^\nu), \qquad (X.\ 10)$$

$$\frac{P_{\nu-1}}{P_\nu}\to\xi_1 \quad (\nu\to\infty). \qquad (X.10a)$$

5. Последнюю формулу можно переписать в виде

$$\xi_1 = \frac{P_0}{P_1} + \sum_{\nu=2}^{\infty} \left(\frac{P_{\nu-1}}{P_\nu} - \frac{P_{\nu-2}}{P_{\nu-1}} \right). \qquad \text{(X. 11)}$$

Представим ν-й член этого ряда, как отношение определителей:

$$\frac{P_{\nu-1}}{P_\nu} - \frac{P_{\nu-2}}{P_{\nu-1}} = -\frac{D_{2,\nu-1}}{D_{1,\nu-1}\, D_{1,\nu}},$$

где

$$D_{2,\nu} = \begin{vmatrix} a_2 & a_1 & \dots & a_{3-\nu} \\ a_3 & a_2 & \dots & a_{4-\nu} \\ \vdots & \vdots & & \vdots \\ a_{\nu+1} & a_\nu & \dots & a_2 \end{vmatrix} \quad (\nu = 1, 2, \dots), \quad D_{2,0} = 1\ ^{1)}.$$

Таким образом, мы получаем следующее разложение для «минимального» корня уравнения (X. 1):

$$\xi_1 = -\sum_{\nu=1}^{\infty} \frac{D_{2,\nu-1}}{D_{1,\nu-1}\, D_{1,\nu}}.$$

---

[1]) Действительно, обозначим через

$$\Delta \begin{pmatrix} \alpha_1, & \alpha_2, & \dots \\ \beta_1, & \beta_2, & \dots \end{pmatrix}$$

определитель, который получен из определителя Δ вычеркиванием в нем строк с индексами $\alpha_1, \alpha_2, \dots$ и столбцов с индексами $\beta_1, \beta_2, \dots$. Тогда по формуле Сильвестра имеем

$$\Delta \begin{pmatrix} 1 \\ 1 \end{pmatrix} \Delta \begin{pmatrix} \nu \\ \nu \end{pmatrix} - \Delta \begin{pmatrix} \nu \\ 1 \end{pmatrix} \Delta \begin{pmatrix} 1 \\ \nu \end{pmatrix} = \Delta\Delta \begin{pmatrix} 1 & \nu \\ 1 & \nu \end{pmatrix}.$$

Возьмем теперь в качестве Δ определитель $D_{1,\nu}$. Тогда

$$\Delta \begin{pmatrix} 1 \\ 1 \end{pmatrix} = D_{1,\nu-1}; \quad \Delta \begin{pmatrix} \nu \\ \nu \end{pmatrix} = D_{1,\nu-1}; \quad \Delta \begin{pmatrix} 1 \\ \nu \end{pmatrix} = D_{2,\nu-1};$$

$$\Delta \begin{pmatrix} \nu \\ 1 \end{pmatrix} = a_0^{\nu-1} = 1; \quad \Delta \begin{pmatrix} 1 & \nu \\ 1 & \nu \end{pmatrix} = D_{1,\nu-2},$$

поэтому из формулы Сильвестра получаем

$$D_{1,\nu-1}^2 - D_{2,\nu-1} = D_{1,\nu}\, D_{1,\nu-2};$$

отсюда и из (X. 6) следует наше утверждение.

Это разложение было исследовано Уиттекером, который доказал его справедливость лишь при весьма специальных предположениях.

Применение полученного ряда требует, однако, совершенно излишних дополнительных вычислений, так как для нахождения $\frac{P_{\nu-1}}{P_\nu}$ с помощью этого ряда необходимо вычислить все определители

$$D_{2,1},\quad D_{2,2},\quad \ldots,\quad D_{2,\nu-1},$$
$$D_{1,1},\quad D_{1,2},\quad \ldots,\quad D_{1,\nu},$$

в то время как для нахождения этого отношения по формуле (X. 6) достаточно вычислить $D_{1,\nu}$ и $D_{1,\nu-1}$.

6. Мы покажем теперь, что знание коэффициентов P разложения (X. 2) позволяет во многих случаях легко вычислить *произведение корней* уравнения (X. 1).

Вместо предположений п. 4 примем условие, что (X.1) имеет внутри своего круга сходимости k корней $\xi_1, \ldots, \xi_k$, таких, что

$$0 < |\xi_1| \leqslant |\xi_2| \leqslant \ldots \leqslant |\xi_k| \tag{X.12}$$

и что модули всех остальных корней этого уравнения *больше*, чем $|\xi_k|$. Положим

$$N(z) = \prod_{\varkappa=1}^{k} (z - \xi_\varkappa) = A_0 z^k + A_1 z^{k-1} + \ldots + A_k, \tag{X. 13}$$
$$A_0 = 1.$$

Тогда можно написать

$$\Phi(z) = \sum_{\nu=0}^{\infty} P_\nu z^\nu = \frac{M(z)}{N(z)} + \sum_{\nu=0}^{\infty} c_\nu z^\nu, \tag{X. 14}$$

где ряд $\sum c_\nu z^\nu$ имеет радиус сходимости $\varrho > |\xi_k|$, а $M(z)$— многочлен степени $< k$, взаимно простой с $N(z)$.

При любом положительном q, удовлетворяющем условию

$$|\xi_k| < \frac{1}{q} < \varrho, \tag{X. 15}$$

имеем

$$c_\nu = o(q^\nu) \quad (\nu \to \infty). \tag{X. 16}$$

Мы получим выражение для произведения $\xi_1 \xi_2 \ldots \xi_k$ через определители

$$\Delta_\nu = \begin{vmatrix} P_\nu & P_{\nu+1} & \ldots & P_{\nu+k-1} \\ P_{\nu-1} & P_\nu & \ldots & P_{\nu+k-2} \\ \vdots & \vdots & \vdots & \vdots \\ P_{\nu-k+1} & P_{\nu-k+2} & \ldots & P_\nu \end{vmatrix}, \quad \nu \geqslant k-1. \qquad (X.\ 17)$$

Наше доказательство опирается только на предположения, содержащиеся в формулах (X. 12), (X. 13) и (X. 14), и не использует формулу (X. 1), т. е. не зависит от предположения о регулярности $\frac{1}{\Phi(z)}$ при $|z|<r$.

7. Разложим $\frac{M(z)}{N(z)}$ по степеням z:

$$\frac{M(z)}{N(z)} = \sum_{\nu=0}^{\infty} y_\nu z^\nu. \qquad (X.\ 18)$$

Умножая обе части этой формулы на $N(z)$ и сравнивая коэффициенты при одинаковых степенях z, получаем уравнение

$$y_\nu A_k + \ldots + y_{\nu-k} A_0 = 0 \qquad (\nu \geqslant k). \qquad (X.\ 19)$$

Рассмотрим теперь определители

$$D_\nu = \begin{vmatrix} y_\nu & y_{\nu+1} & \ldots & y_{\nu+k-1} \\ y_{\nu-1} & y_\nu & \ldots & y_{\nu+k-2} \\ \vdots & \vdots & \vdots & \vdots \\ y_{\nu-k+1} & y_{\nu-k+2} & \ldots & y_\nu \end{vmatrix}, \quad \nu \geqslant k-1. \qquad (X.\ 20)$$

Прибавим к первой строке определителя (X.20) вторую строку, умноженную на $\frac{A_{k-1}}{A_k}$, третью строку, умноженную на $\frac{A_{k-2}}{A_k}$, ..., последнюю строку, умноженную на $\frac{A_1}{A_k}$; тогда первая строка вследствие формулы (X.19) перейдет в строку

$$-\frac{y_{\nu-k}}{A_k}, \quad -\frac{y_{\nu-k+1}}{A_k}, \ldots, \quad -\frac{y_{\nu-1}}{A_k}.$$

Если с помощью $(k-1)$-й перестановки сделать эту строку последней, то получим определитель, отличающийся от $D_{\nu-1}$ только множителем

$$\frac{(-1)^k}{A_k} = \frac{1}{\xi_1 \ldots \xi_k}.$$

Таким образом,

$$D_{\nu-1} = \xi_1 \ldots \xi_k D_\nu. \tag{X. 21}$$

8. Докажем теперь, что определитель

$$D_{k-1} = \begin{vmatrix} y_{k-1} & y_k & \ldots & y_{2k-2} \\ y_{k-2} & y_{k-1} & \ldots & y_{2k-3} \\ \vdots & \vdots & \vdots & \vdots \\ y_0 & y_1 & \ldots & y_{k-1} \end{vmatrix}$$

не равен нулю. Действительно, допустим противное. Тогда из равенства $D_{k-1} = 0$ следует, что система линейных уравнений

$$\begin{cases} \alpha_{k-1} \quad y_0 + \alpha_{k-2} \quad y_1 + \ldots + \alpha_0 \; y_{k-1} = 0, \\ \alpha_{k-1} \quad y_1 + \alpha_{k-2} \quad y_2 + \ldots + \alpha_0 \; y_k = 0, \\ \vdots \qquad\qquad \vdots \qquad\qquad\qquad \vdots \\ \alpha_{k-1} \; y_{k-1} + \alpha_{k-2} \; y_k + \ldots + \alpha_0 \; y_{2k-2} = 0 \end{cases} \tag{X. 22}$$

имеет ненулевое решение, т. е. что существуют k чисел $\alpha_0, \alpha_1, \ldots, \alpha_{k-1}$, которые не все равны нулю и которые удовлетворяют системе (X. 22).

Умножим обе части равенства (X. 18) на многочлен

$$H(z) \equiv \alpha_{k-1} z^{k-1} + \alpha_{k-2} z^{k-2} + \ldots + \alpha_0.$$

В произведении $H(z) \sum_{\nu=0}^{\infty} y_\nu z^\nu$ все коэффициенты при $z^{k-1}, z^k, \ldots, z^{2k-2}$ равны нулю вследствие (X. 22). Поэтому

$$H(z) \cdot \frac{M(z)}{N(z)} = T_{k-2}(z) + z^{2k-1} W(z),$$

где $T_{k-2}(z)$ есть многочлен степени $k-2$, а $W(z)$ — степенной ряд, содержащий только неотрицательные степени z. Но тогда из равенства

$$\frac{H(z)\dfrac{M(z)}{N(z)}-T_{k-2}(z)}{z^{2k-1}}=W(z)$$

следует, что многочлен $H(z)M(z)-T_{k-2}(z)N(z)$, имеющий степень не более чем $2k-2$, делится на z^{2k-1}; поэтому он должен быть тождественно равен нулю. Таким образом, мы приходим к равенству

$$\frac{M(z)}{N(z)}=\frac{T_{k-2}(z)}{H(z)},$$

которое противоречит условию, что $M(z)$ и $N(z)$ взаимно просты. Мы видим, что определитель D_{k-1} не может быть равен нулю[1]).

9. Из (X.14) и (X.18) получаем

$$P_\nu=y_\nu+c_\nu \quad (\nu=0,\ 1,\ldots). \tag{X. 23}$$

Поэтому можно ожидать, что определитель Δ_ν мало отличается от D_ν. Однако для получения соответствующей оценки мы должны будем исследовать матрицу, обратную к матрице[2])$(D\ )$.

Возьмем матрицу

$$X=\begin{pmatrix} -A_1 & 1 & 0 & \ldots & 0 \\ -A_2 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ -A_{k-1} & 0 & 0 & \ldots & 1 \\ -A_k & 0 & 0 & \ldots & 0 \end{pmatrix}, \tag{X. 24}$$

где $A_1, A_2, \ldots, A_k$ — коэффициенты многочлена $N(z)$ в (X. 13), и рассмотрим произведение $(D_\nu)X$. Чтобы получить первый

---

[1]) Этот результат был, по-видимому, впервые получен Кронекером в 1881 г.

[2]) Символом (D) мы будем обозначать матрицу, соответствующую определителю D.

элемент первой строки этого произведения, умножим первую строку матрицы (D_ν)

$$y_\nu,\ y_{\nu+1},\ \dots,\ y_{\nu+k-1} \tag{X. 25}$$

на первый столбец матрицы X; мы получим в силу (X. 19)

$$-A_1 y_\nu - A_2 y_{\nu+1} - \dots - A_k y_{\nu+k-1} = y_{\nu-1}.$$

Следующие элементы первой строки произведения $(D_\nu) X$ получаются умножением (X. 25) на второй, третий, ..., k-й столбцы матрицы X и поэтому равны $y_\nu,\ y_{\nu+1}, \dots,\ y_{\nu+k-2}$. Таким образом, первая строка произведения $(D_\nu) X$ есть

$$y_{\nu-1},\ y_\nu,\ y_{\nu+1}, \dots,\ y_{\nu+k-2}. \tag{X. 26}$$

Остальные строки произведения получаются из (X. 26) последовательным уменьшением индексов в (X.26) на единицу, так как тем же свойством обладают строки матрицы (D_ν).

Мы видим, что произведение $(D_\nu) X$ равно матрице $(D_{\nu-1})$:

$$(D_\nu) X = (D_{\nu-1}).$$

Отсюда следует, что

$$(D_\nu) X^{\nu-k+1} = (D_{k-1}); \quad (D_\nu)^{-1} = X^{\nu-k+1} (D_{k-1})^{-1},$$

или

$$(D_\nu)^{-1} = X^\nu S, \tag{X. 27}$$

где

$$S = X^{-k+1} (D_{k-1})^{-1}.$$

10. Положим

$$C_\nu = \begin{pmatrix} c_\nu & c_{\nu+1} & \dots & c_{\nu+k-1} \\ c_{\nu-1} & c_\nu & \dots & c_{\nu+k-2} \\ \vdots & \vdots & \vdots & \vdots \\ c_{\nu-k+1} & c_{\nu-k+2} & \dots & c_\nu \end{pmatrix}. \tag{X. 28}$$

Из (X. 17), (X. 20), (X. 23) и (X. 27) имеем

$$(\Delta_\nu) = (D_\nu) + C_\nu,$$

$$(\Delta_\nu)(D_\nu)^{-1} = E + C_\nu X^\nu S.$$

Переходя к определителям, получаем

$$\frac{\Delta_\nu}{D_\nu} = |E + C_\nu X^\nu S|. \tag{X.29}$$

Теперь мы должны получить оценки для X^ν при $\nu \to \infty$. Для этого нам понадобятся собственные значения матрицы X.

Докажем, что

$$|zE - X| = \begin{vmatrix} A_1+z & -1 & 0 & \dots & 0 & 0 & 0 \\ A_2 & z & -1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ A_{k-2} & 0 & 0 & \dots & z & -1 & 0 \\ A_{k-1} & 0 & 0 & \dots & 0 & z & -1 \\ A_k & 0 & 0 & \dots & 0 & 0 & z \end{vmatrix} = N(z). \tag{X. 30}$$

Действительно, полагая при $\nu \leqslant k$,

$$N_\nu(z) = A_0 z^\nu + \dots + A_\nu, \quad N_k(z) = N(z),$$

имеем

$$N_\nu(z) = zN_{\nu-1}(z) + A_\nu.$$

При $k=1$ равенство (X. 30) справедливо:

$$N_1(z) = z + A_1 = A_0 z + A_1.$$

Предположим теперь, что соотношение типа (X. 30) справедливо при $\nu \leqslant k-1$, и разложим определитель (X. 30) по элементам последней строки. Мы получим

$$|zE - X| = (-1)^{k-1} A_k (-1)^{k-1} + zN_{k-1}(z) = N(z).$$

Таким образом, собственными значениями матрицы X являются [1])

$$\xi_1, \xi_2, \dots, \xi_k.$$

В частности, $\lambda_x = |\xi_k|$ (см. гл. 15, п. 6).

11. Оценим теперь матрицу X^ν с помощью теоремы 16.1, полагая все U_μ равными нулю. При любом $\varepsilon > 0$ и соответствующем $\sigma = \sigma(X, \varepsilon) > 0$ получаем

$$|X^\nu|_\infty < \sigma(\lambda_X + \varepsilon)^\nu = \sigma(|\xi_k| + \varepsilon)^\nu \quad (\nu = 1, 2, \dots).$$

---

[1]) Этот результат принадлежит С. Гюнтеру (1876).

С другой стороны, из (X. 16) получаем оценку нормы матрицы (X. 28):

$$|C_\nu|_\infty = o(q^\nu) \quad (\nu \to \infty).$$

В силу (15. 19) поэтому имеем

$$|C_\nu X^\nu|_\infty = o[(q|\xi_k| + q\varepsilon)^\nu],$$
$$|C_\nu X^\nu S|_\infty = o[(q|\xi_k| + q\varepsilon)^\nu] \quad (\nu \to \infty).$$

Таким образом, полагая

$$q|\xi_k| + q\varepsilon = \vartheta,$$

получаем из формулы (X. 29) оценку

$$\frac{\Delta_\nu}{D_\nu} = 1 + o(\vartheta^\nu) \quad (\nu \to \infty).$$

В этой оценке ϑ может быть любым числом, удовлетворяющим неравенству

$$\frac{|\xi_k|}{\rho} < \vartheta < 1.$$

Действительно, если, взяв любое такое ϑ, положим

$$\delta = \vartheta - \frac{|\xi_k|}{\rho}; \quad q = \frac{1}{\rho} + \frac{\delta}{2|\xi_k|}; \quad \varepsilon = \frac{\delta}{2q},$$

то будут выполнены как неравенство (X. 15), так и равенство

$$\vartheta = q|\xi_k| + q\varepsilon.$$

12. Наконец, из формулы (X. 21) получаем при любом ϑ, $\frac{|\xi_k|}{\rho} < \vartheta < 1$:

$$\frac{\Delta_{\nu-1}}{\Delta_\nu} = \xi_1 \ldots \xi_k + o(\vartheta^\nu); \qquad \frac{\Delta_{\nu-1}}{\Delta_\nu} \to \xi_1 \ldots \xi_k. \qquad (X.31)$$

При выводе соотношений (X. 31) мы предполагали только, что $\Phi(z)$ мероморфна в круге $|z| < \varrho$ и внутри этого круга не имеет других полюсов, кроме точек $\xi_1, \ldots, \xi_k$. Если же $\Phi(z)$ задана формулой (X. 2), то мы должны предположить, что $f(z)$ в круге $|z| < \varrho$ обращается в нуль в точках $\xi_1, \ldots, \xi_k$, а все остальные нули этой функции, если они существуют, имеют модули, большие, чем

$$\max(|\xi_1|, \ldots, |\xi_k|).$$

Если, в частности, $f(z)$ — многочлен, то формула (X. 31) дает обобщение так называемого метода Бернулли, в котором рассматривается только случай $k=1$. Линейный характер сходимости метода Бернулли хорошо известен.

Первая из формул (X. 31) может быть легко получена из известной теоремы Адамара (J. Hadamard, Essai sur l'étude des fonctions données par leur développement de Taylor, *J. math. pures et appl.* (4 série), 8 (1892), 101—186).

13. Если в (X. 12) имеют место строгие неравенства:

$$|\xi_1|<|\xi_2|<|\xi_3|<\ldots<|\xi_k|, \tag{X.32}$$

то корни $\xi_1, \ldots, \xi_k$ можно найти последовательно, не вычисляя определителей Δ_ν. Действительно, найдем ξ_1 по формуле (X. 10a); умножая затем (X. 2) на $1-\frac{z}{\xi_1}$, получаем

$$\frac{1-\frac{z}{\xi_1}}{f(z)}=\sum_{\nu=0}^{\infty} Q_\nu z^\nu, \qquad Q_0=1, \tag{X. 33}$$

$$Q_\nu=P_\nu-\frac{1}{\xi_1}P_{\nu-1} \quad (\nu>0).$$

Отсюда и из формулы (X. 10) получаем

$$\frac{Q_{\nu-1}}{Q_\nu}\to\xi_2. \tag{X.34}$$

Умножая теперь ряд (X. 33) на $1-\frac{z}{\xi_2}$, получаем ряд

$$\sum_{\nu=0}^{\infty} R_\nu z^\nu, \quad R_0=1; \quad R_\nu=Q_\nu-\frac{1}{\xi_2}Q_{\nu-1} \quad (\nu>0), \tag{X. 35}$$

у которого

$$\frac{R_{\nu-1}}{R_\nu}\to\xi_3, \tag{X. 36}$$

и т. д. Если с самого начала коэффициенты P_ν в (X. 2) вычислены вплоть до P_n, то в (X. 33) можно заменить

$\frac{1}{\xi_1}$ на $\frac{P_n}{P_{n-1}}$, что дает следующую аппроксимацию корня ξ_2:

$$\frac{Q_{n-2}}{Q_{n-1}} = \frac{P_{n-1}P_{n-2} - P_n P_{n-3}}{P_{n-1}^2 - P_n P_{n-2}}. \qquad (X.\ 37)$$

Заменяя теперь ξ_2 приближением (X. 37), получаем в качестве приближенного значения ξ_3 отнсшение $\frac{R_{n-3}}{R_{n-2}}$, и т. д. Однако этот метод применим только в том случае, когда выполнено условие (X. 32).

14. Рассмотрим в качестве примера функцию

$$\cos\sqrt{x} = \sum_{\nu=0}^{\infty}(-1)^\nu \frac{x^\nu}{(2\nu)!},$$

у которой

$$\xi_1 = \frac{\pi^2}{4} = 2{,}46740, \quad \xi_1\xi_2 = 9\frac{\pi^4}{16} = 54{,}79.$$

В этом случае коэффициенты P_ν с $\nu = 0, 1, \ldots, 5$ имеют следующие значения:

ν	0	1	2	3	4	5
P_ν	1	0,5	0,208333	0,08472222	0,034350198	0,013922233

Ниже приведена таблица последовательных приближений $\frac{P_{\nu-1}}{P_\nu}$ (с соответствующими погрешностями E_ν):

ν	1	2	3	4	5
$\frac{P_{\nu-1}}{P_\nu}$	2	2,4	2,4590	2,466426	2,467291
E_ν	0,47	0,067	0,0084	0,0010	0,00011

Далее, мы получаем для $k=2$ и $\nu=2, 3, 4$ величины $\frac{\Delta_{\nu-1}}{\Delta_\nu}$, аппроксимирующие произведение $\xi_1\xi_2$ с погрешностями E'_ν

ν	2	3	4
$\frac{\Delta_{\nu-1}}{\Delta_\nu}$	40,01	48,19	52,25
E'_ν	14,78	6,60	2,54

В этом случае погрешности E_ν и E'_ν действительно убывают и отношение $\frac{E_\nu}{E_{\nu-1}}$ стремится к $\frac{1}{9}=\frac{\pi^2}{4}\frac{1}{(3\pi/2)^2}$, а отношение $\frac{E'_\nu}{E'_{\nu-1}}$ стремится к $\frac{9}{25}=\frac{(3\pi/2)^2}{(5\pi/2)^2}$.

ПРИЛОЖЕНИЕ XI

НЕПРЕРЫВНОСТЬ СОБСТВЕННЫХ ЗНАЧЕНИЙ КАК ФУНКЦИЙ ОТ ЭЛЕМЕНТОВ МАТРИЦЫ

1. Теорема. Пусть $A = (a_{\mu\nu})$, $B = (b_{\mu\nu})$ — две квадратные матрицы порядка n, и пусть

$$\varphi(\lambda) \equiv |A - \lambda E| = 0, \quad \psi(\lambda) \equiv |B - \lambda E| = 0 \qquad \text{(XI. 1)}$$

— соответствующие характеристические многочлены и уравнения. Обозначим корни многочлена $\varphi(\lambda)$ через λ_ν, а корни многочлена $\psi(\lambda)$ — через λ'_ν. Положим

$$M = \max(|a_{\mu\nu}|, |b_{\mu\nu}|) \quad (\mu, \nu = 1, \ldots, n), \qquad \text{(XI. 2)}$$

$$\frac{1}{nM} \sum_{\mu, \nu} |a_{\mu\nu} - b_{\mu\nu}| = \delta. \qquad \text{(XI. 3)}$$

Тогда любому корню λ'_ν многочлена $\psi(\lambda)$ можно поставить в соответствие такой корень λ_ν многочлена $\varphi(\lambda)$, чтобы выполнялось неравенство

$$|\lambda'_\nu - \lambda_\nu| \leqslant (n+2) M\delta^{1/n}. \qquad \text{(XI. 4)}$$

Далее, можно так перенумеровать λ_ν и λ'_ν, чтобы при любом ν иметь

$$|\lambda'_\nu - \lambda_\nu| \leqslant 2(n+1)^2 M\delta^{1/n} \quad (\nu = 1, \ldots, n). \qquad \text{(XI. 5)}$$

2. Докажем сначала следующую лемму.

Лемма. Если выполнены условия теоремы, то при любом λ, у которого $|\lambda| \leqslant nM$,

$$|\varphi(\lambda) - \psi(\lambda)| \leqslant (n+2)^n M^n \delta. \qquad \text{(XI. 6)}$$

Доказательство. Расположим элементы матрицы A в произвольном порядке и обозначим их $\alpha_1, \ldots, \alpha_{n^2}$; *соответствующие элементы* матрицы B обозначим $\beta_1, \ldots, \beta_{n^2}$. Тогда при фиксированном λ можно написать

$$\varphi(\lambda) = P(\alpha_1, \ldots, \alpha_{n^2}), \quad \psi(\lambda) = P(\beta_1, \ldots, \beta_{n^2}),$$

$$\varphi(\lambda) - \psi(\lambda) = \sum_{\varkappa=1}^{n^2} \Delta_\varkappa, \tag{XI. 7}$$

где

$$\Delta_\varkappa = P(\alpha_1, \ldots, \alpha_\varkappa, \beta_{\varkappa+1}, \ldots, \beta_{n^2}) - P(\alpha_1, \ldots, \alpha_{\varkappa-1}, \beta_\varkappa, \ldots, \beta_{n^2}).$$

Так как многочлен $P(\alpha_1, \ldots, \alpha_{n^2})$ линеен по всем аргументам α_ν, то

$$\Delta_\varkappa = \pm(\alpha_\varkappa - \beta_\varkappa) T_\varkappa, \tag{XI. 8}$$

где $T_\varkappa$ получается из минора определителя P, имеющего порядок $n-1$, заменой в нем некоторых α_i на соответствующие им β_i. Далее, так как в каждой строке минора $T_\varkappa$ разность между λ и каким-нибудь из элементов α_j, β_j встречается не более одного раза, то норма[1]) каждой строки в $T_\varkappa$ не превосходит

$$\sqrt{(n-2)M^2 + (n+1)^2 M^2} < (n+2)M.$$

Воспользовавшись оценкой для модуля определителя, данной Адамаром, получаем[2])

$$|T_\varkappa| \leqslant (n+2)^{n-1} M^{n-1}.$$

Отсюда и из формул (XI. 3), (XI. 7), (XI. 8) получаем, наконец,

$$\varphi|(\lambda) - \psi(\lambda)| \leqslant (n+2)^{n-1} M^{n-1} \sum_{\varkappa=1}^{n^2} |\alpha_\varkappa - \beta_\varkappa| \leqslant (n+2)^n M^n \delta.$$

Лемма доказана.

---

[1]) Здесь в качестве нормы строки $\xi_\mu = (c_{\mu,1}, \ldots, c_{\mu,n-1})$ берется число $\xi_\mu| = \sqrt{\sum_{\nu=1}^{n-1} |c_{\mu\nu}|^2}$. — *Прим. перев.*

[2]) Это неравенство можно получить, например, из следующего: если $C = (c_{\mu\nu})$, то $|C| \leqslant \prod_{\mu=1}^{n-1} |\xi_\mu|$. — *Прим. перев.*

3. Перейдем к доказательству теоремы. Для любого корня λ' многочлена $\psi(\lambda)$ имеем по теореме 15.1 неравенство $|\lambda'| \leqslant nM$, и, следовательно, по (XI. 6)

$$|\varphi(\lambda')| = |\varphi(\lambda') - \psi(\lambda')| \leqslant (n+2)^n M^n \delta,$$

$$\prod_{\nu=1}^{n} |\lambda' - \lambda_\nu| \leqslant [(n+2) M\delta^{1/n}]^n; \qquad \text{(XI. 9)}$$

поэтому по крайней мере для одного λ_ν выполнено неравенство (XI. 4). Если мы обозначим теперь правую часть неравенства (XI. 4) через ε, то рассуждения, проведенные в п. 6, 7 приложения I, могут быть без изменений повторены в данном случае. Поэтому, перенумеровав подходящим образом λ_ν и λ'_ν, получим

$$|\lambda'_\nu - \lambda_\nu| \leqslant 2n(n+2)M\delta^{1/n} \leqslant 2(n+1)^2 M\delta^{1/n} \quad (\nu = 1, \ldots, n).$$

Теорема полностью доказана.

БИБЛИОГРАФИЧЕСКИЕ ПРИМЕЧАНИЯ

СТАТЬИ И КНИГИ, НА КОТОРЫЕ БЫЛИ НЕОДНОКРАТНЫЕ ССЫЛКИ

1. Householder A. S., Principles of numerical analysis, McGraw-Hill, New York, 1953 (русский перевод: Хаусхолдер А. С., Основы численного анализа, М., ИЛ, 1956).
2. Ostrowski A., Vorlesungen über Differentia-lund Integralrechnung, vol. II, Birkhäuser, Basel, 1951.
3. Островский А., О сходимости и оценках погрешностей некоторых методов решения численных уравнений, «Сборник, посвященный памяти академика Дмитрия Александровича Граве», стр. 213—234, М.—Л., Гостехиздат, 1940.
4. Ostrowski A., Recherches sur la méthode de Gräffe et les zéros des polinômes et des séries de Laurent, *Acta Math.*, **72** (1940), 99—257.
5. Schröder E., Über unendlich viele Algorithmen zur Auflösung der Gleichungen, *Math. Ann.*, **2** (1870), 317—365.
6. Steffensen J. F., Interpolation, Chelsea Publ., New York, 1950.
7. Willers F. A., Methoden der praktischen Analysis (2-е издание), de Gruyter, Berlin, 1954.

Глава 1

Изложение частично заимствовано из книги Островского [2], стр. 184, 298.

Глава 2

Замечания об обратной интерполяции см. у Стеффенсена [6], стр. 80. Теорема Дарбу впервые дана в фундаментальной статье Дарбу: Mémoire sur les fonctions discontinues, *Ann. Ecole Norm. Sup. Paris*, **4**, 1875, а формулы (2.5) и (2.7) в статье Шредера [5], стр. 330. Теорема 2.2, по-видимому, нигде не опубликована.

Глава 3

Метод regula falsi восходит к первым итальянским алгебраистам. Материал, изложенный в п. 4, 5 и 9—16, ранее не опубликован.

Глава 4

Различение точек притяжения и точек отталкивания было введено позднее Дж. Р. Риттом. Лемма 4.2 для случая точек притяжения дана еще Шредером [5], стр. 323. Материал п. 6—8, по-видимому, нигде не опубликован.

Глава 5

Теорема 5.1 является улучшенным вариантом теоремы Мизеса и Гейрингер, данной ими в статье Mises R., Geiringer H., Praktische Verfahren der Gleichungsauflösung, zusammenfassender Bericht, *Z. angew. Math. Mech.*, **9** (1929), 58—77.

Некоторые важные пункты исследования Мизеса и Гейрингер восходят к Бауэру и Фейеру. Ср. с работой Bauer M., Zur Bestimmung der reelen Wurzeln einer algebraischen Gleichung durch Iteration, *Jahresber. deut. Math. — Ver.*, **25** (1916), 294—301. Теорема 5.2 впервые опубликована Островским [3] и позднее распространена на более широкий класс сравниваемых функций в статье Карамата: Karamata J., Über das asimptotische Verhalten der Folgen, die durch Iteration definiert sind, *Rec. trav. Acad. Serbe Sci.*, **35** (1953), 60. Материал п. 7 ранее не был опубликован.

Глава 6

Квадратичный порядок сходимости исследован Ж. Б. Фурье. См. Oeuvres de Fourier, т. II, 249—250, Gauthier—Villars, Paris, 1890.

Глава 7

Об оценках a priori и a posteriori см. Ostrowski A., Sur la continuité relative des racines d'équations, *Compt. rend.*, **209** (1939), 777—779.

Теоремы существования в данной форме впервые опубликованы Островским [3]. В несколько менее строгой форме эти теоремы были даны Коши в его Leçons sur le calcul differentiel, Paris, Buré frères, 1829, перепечатанных в Oeuvres complètes, 2-me ser., t. IV, Gauthier—Villars, Paris, 1899; см., в частности, стр. 576, 578, 593—594. Что касается теоремы 7.1, то Коши требовал вместо условия

$$|f'(z_0)| \geqslant 2\,|h_0|\,M$$

выполнения условия

$$\min_{K_0} |f'(z)| > 4\,|h_0|\,M.$$

В теореме 7.2 условие Коши имело вид: $\min\limits_{J_0} |f'(x)| > 2\,|h_0|\,M$ вместо (7.1).

Глава 8

Формула (8.1) принадлежит Шредеру [5], стр. 325. Теорема 8.1 ранее не была опубликована.

Глава 9

Границы (9.1) даны впервые Ж.Б. Фурье в 1818 г. (см. Oeuvres de Fourier, t. II, 248, Gauthier — Villars, Paris, 1890).

Исследование, проведенное в гл. 9, и, в частности, формулы (9.3), (9.15), (9.16), ранее нигде не были опубликованы.

Глава 10

Границы (10.1) предложены Г. П. Данделеном в 1824 г.; см. *N. Mém. Acad. Bruxelles*, **3** (1826).

Теорема 10.1 нигде ранее не была опубликована.

Глава 11

Материал этой главы нигде ранее не был опубликован.

Глава 12

Приведенное здесь изложение общей теории разностных уравнений с постоянными коэффициентами ср., например, с изложением в книге Milne-Thomson L. M., The Calculus of Finite Differences, стр. 384—414, Mac-Millan, London, 1933 или Nörlund N. E., Differenzenrechnung, 295—300, Springer, Berlin, 1924. Некоторые результаты этой главы, по-видимому, нигде не опубликованы.

Глава 13

Ее содержание ранее не было опубликовано. Теорема Энестрема и Какейа была приведена в статье Энестрема (написанной на шведском языке) Eneström, Härledning af en allmän formel för antalet pensionärer, som vid en godtycklig tidpunkt förefinnas inom en sluten pensionskassa, *Ofversigt af vetenskaps, Acad. Förhandl.*, **50** (1893), 405—415 и оставалась, по-видимому, совершенно неизвестной. Она была вновь получена Какейа и опубликована в статье Kakeya S., On the limits of the roots of an algebraic equation with positive coefficients, *Tôhoku Math. J.* **2** (1912), 140.

Глава 14

Бесконечный ряд (14.2) был впервые рассмотрен Ньютоном и Эйлером; этот ряд был изучен также (без исследования сходимости) Теремином (Theremin, *J. f. reine angew. Math.*, **49** (1855), 178—243) и Шредером [5], стр. 329. Теоремы 14.1 и 14.2 нигде не опубликованы.

Глава 15

Результаты, относящиеся к нормам векторов и матриц, соответствующим $p = 1, \infty$, даны в явной форме В. Н. Фаддеевой в книге «Вычислительные методы линейной алгебры», Москва, Физматгиз, 1950 г. Более ранние формулировки можно найти в статьях Rella T Über den absoluten Betrag von Matrizen, Proc. Intern. Congr. Math., Oslo, 1936, t. II, 29—31 и Über positiv-homogene Funktionen ersten Grades einer Matrix, *Monatsh. Math. u. Physik*, **48** (1939), 84—95. Более общее исследование норм матриц см. в статьях Ostrowski A., Über Normen von Matrizen, *Math. Z.*, **63**, 1955, 2—18 и Householder A. S., On norms of vectors and matrices, Oak Ridge Natl. Lab. Rept., **1756**, 1954. Теорема 15.1 принадлежит Фробениусу.

Теорема 15.3 нигде не опубликована.

Главы 16 и 17

Результаты были опубликованы автором без доказательства в *Comptes rendus*, **244** (1957), 288—289.

Глава 18

Теоремы 18.1 и 18.2 были опубликованы автором без доказательств в *Comptes rendus*, **244** (1957), 288—289. Критерий (18.23) сходимости итерации принадлежит Скарборо: Scarborough A. D., Numerical Mathematical Analysis, 1 ed., Johns Hopkins Press, Baltimore, 1930 (русский перевод: Скарборо А. Д., Численные методы математического анализа, М.—Л., ГИТТЛ, 1934).

Другой специальный случай рассмотрен в статье Schulz G., Über die Lösung von Gleichungen durch Iteration, *Z. angew. Mat. Mech.*, **22** (1942), 234—235.

Приложение I

Эти результаты впервые получены Островским [4], стр. 209—212, 218.

Приложение II

Более точные результаты получены Островским [4], стр. 212—217, однако в данной книге доказательство более элементарно.

Приложение III

Опубликовано без доказательств в 1957 г., *Compt. rend.*, **244**, 429—430.

Приложение IV

Изложение только по форме отличается от цитированной книги Гаусса.

Приложение V

Библиография дана в примечаниях к тексту.

Приложение VI

Ранее не опубликовано.

Приложение VII

Результаты п. 1—3 содержатся в статье Островского [3]. Материал п. 4—9 нигде не опубликован.

Приложение VIII

Ранее не опубликовано.

Приложение IX

Здесь улучшены в некоторых направлениях результаты, опубликованные в статье Ostrowski A. M., A Method of speeding up iterations with super-linear convergence, *J. Math. Mech.*, **7** (1958), 117—120.

Приложение X

Исследования этого приложения вызваны желанием автора выяснить обоснование разложения Уиттекера (п. 5), опубликованного в его статье Whittaker E. T., A formula for the solution of algebraic or transcendental equations, *Proc. Math. Soc. Edinburgh*, **36** (1918), 103—106, и в книге Whittaker E. T. and Robinson G., The Calculus of Observation, Blackie and Son, London and Glasgow, 1928, § 60 (русский перевод: Уиттекер Э. Т. и Робинсон Д., Математическая обработка результатов наблюдений, М., ОНТИ, 1935).

Идея этой аппроксимации принадлежит де Моргану А.; см. *J. Inst. Actuaries*, **14** (1868), 353. В остальном содержание этого приложения, по-видимому, ранее нигде не было опубликовано.

Приложение XI

Результаты были впервые опубликованы в статье Ostrowski A., Über die Stetigkeit von charakteristischen Wurzeln in Abhängigkeit von den Matrizenelementen, *Jahresber. deut. Mat. — Ver.*, **60** (1957), 40—42.

УКАЗАТЕЛЬ

ОГЛАВЛЕНИЕ

Приложения

А. М. Островский

РЕШЕНИЕ УРАВНЕНИЙ
И СИСТЕМ УРАВНЕНИЙ

Редактор Г. К. Москатов
Художник М. Л. Компанеец
Технический редактор А. Д. Хомяков
Корректор *Т. С. Бухтина*
Сдано в производство 15/XI 1962 г.
Подписано к печати 20/IV 1963 г.
Бумага $84 \times 108^1/_{32} = 3,4$, бум. л.
11,3 печ. л.
Уч.-изд. л. 8,9. Изд. № 1/1335
Цена 77 к. Зак. 825

ИЗДАТЕЛЬСТВО
ИНОСТРАННОЙ ЛИТЕРАТУРЫ
Москва, 1-й Рижский пер., 2

1-я тип. Трансжелдориздата МПС
Б. Переяславская, д. 46