Численный анализ

А.М. Мацокин, ВКИ НГУ, 1994/95 учебный год
(конспект лекций)

Содержание

1. Алгебраические методы интерполирования

Задача интерполирования функции $f(x)$ состоит в том, чтобы по известным ее значениям в некоторых точках определить ее значения в остальных точках области задания. Такая задача возникает, например, тогда, когда по результатам измерения некоторой физической величины в одних точках требуется определить ее значения в других точках или когда в целях ускорения вычислений желательно приблизить заданную функцию другой, но "легко" вычислимой. Поскольку наиболее простыми с вычислительной точки зрения являются алгебраические многочлены (полиномы), то вполне естественным является их использование при решении интерполяционных задач.

Определение 1. Алгебраический полином

$$P_m(x) = a_0 + a_1 \cdot x + ... + a_m \cdot x^m \tag{1}$$

называется интерполяционным полиномом для функции $f(x)$, заданной на отрезке $[\,a,\ b\,]$, по ее значениям $f(x_i)$ в $n+1$ попарно различных точках (узлах) $x_i \in [\,a,\ b\,]$, $i = 0, 1, ..., n$, если

$$P_m(x_i) = f(x_i). \tag{2}$$

Для определения полинома $P_m(x)$ ($m+1$ неизвестное: a_0, $a_1, ...,$ a_m) мы имеем $n+1$ условие (уравнения (2)). Единственность решения математических конечномерных задач обычно обеспечивается равенством числа неизвестных количеству накладываемых на них условий. Если условий меньше, чем неизвестных, то решение обычно не единственно; если условий больше, то решения вообще может не существовать.

Теорема 1. *Задача алгебраической интерполяции (2) при $m = n$ имеет единственное решение.*

Доказательство. Перепишем систему (2) относительно неизвестных коэффициентов a_0, a_1, ... , a_n полинома$P_n(x)$ в развернутом виде:

$$\begin{array}{c} a_0 + x_0 \cdot a_1 + ... + x_0^n \cdot a_n = f(x_0), \\ a_0 + x_1 \cdot a_1 + ... + x_1^n \cdot a_n = f(x_1), \\ \dots\dots\dots \\ a_0 + x_n \cdot a_1 + ... + x_n^n \cdot a_n = f(x_n). \end{array} \tag{3}$$

Определитель матрицы этой системы линейных алгебраических уравнений является известным определителем Вандермонда и отличен от нуля, так как $x_i \neq x_j$ при $i \neq j$, $i, j = 0, ..., n$, что является необходимым и достаточным условием существования и единственности решения системы (3). $\square$

1.1. Интерполяционный полином в форме Лагранжа

Теорема 2. *Решение задачи алгебраической интерполяции (2) при $m = n$ представимо в форме Лагранжа*

$$P_n(x) = \sum_{k=0}^{n} \frac{\omega(x)}{(x - x_k) \cdot \omega'(x_k)} f(x_k)\,, \tag{4}$$

где

$$\omega(x) = (x - x_0) \cdot (x - x_1) \cdot \ldots \cdot (x - x_n) \tag{5}$$

– полином степени $n + 1$.

Доказательство. Для каждого $k = 0, 1, \ldots, n$ рассмотрим частный случай задачи алгебраического интерполирования:

$$P_{n,k}(x_i) = \delta_{k,i}, \quad i = 0, 1, \ldots, n\,,$$

где $\delta_{k,i} = 0$ при $k \neq i$, $\delta_{k,k} = 1$.

Так как полином $P_{n,k}(x)$ степени n по условию имеет n корней

$$x_0\,,\ x_1\,,\ \ldots,\ x_{k-1}\,,\ x_{k+1}\,,\ \ldots,\ x_n\,,$$

то он может быть представлен в виде произведения мономов:

$$P_{n,k}(x) = q_k \cdot (x - x_0) \cdot (x - x_1) \cdot \ldots \cdot (x - x_{k-1}) \cdot (x - x_{k+1}) \cdot \ldots \cdot (x - x_n)\,, \tag{6}$$

где q_k определяется из условия $P_{n,k}(x_k) = 1$:

$$q_k = \frac{1}{(x_k - x_0) \cdot (x_k - x_1) \cdot \ldots \cdot (x_k - x_{k-1}) \cdot (x_k - x_{k+1}) \cdot \ldots \cdot (x_k - x_n)}\,.$$

Легко проверить, что этот полином можно переписать в следующем виде:

$$P_{n,k}(x) = \frac{\omega(x)}{(x - x_k) \cdot \omega'(x_k)}\,, \tag{7}$$

где $\omega(x) = (x - x_0) \cdot (x - x_1) \cdot \ldots \cdot (x - x_n)$.

Очевидно, что линейная комбинация полиномов $P_{n,k}(x)$ с весами $f(x_k)$ является полиномом степени n и, так как

$$\sum_{k=0}^{n} P_{n,k}(x_i) \cdot f(x_k) = \sum_{k=0}^{n} \delta_{k,i} \cdot f(x_k) = f(x_i)$$

для $i = 0, 1, \ldots, n$, она является интерполянтом функции $f(x)$.

1.2. Интерполяционный полином в форме Ньютона

Рассмотрим следующую ситуацию. Предположим, что мы построили интерполяционный полином $P_n(x)$ и вычислили его значения для некоторого набора точек отрезка $[\,a,\ b\,]$, но появились новые значения интерполируемой функции $f(x)$, скажем одно в новой точке x_{n+1}. Можно ли подправить уже вычисленную таблицу значений, т.е. добавить к ней поправку $P_{n+1}(x) - P_n(x)$?

Пусть $P_k(x)$ интерполяционный полином для функции $f(x)$, заданной на отрезке $[\,a,\ b]$, по ее значениям $f(x_i)$ в первых $k+1$ узле из $n+1$ попарно различных точек $x_0,\, x_1,\, \dots,\, x_n$.

Так как разность $P_{k+1}(x) - P_k(x)$ является полиномом степени $k+1$, а значения интерполяционных полиномов $P_{k+1}(x)$ и $P_k(x)$ при $x = x_0, x_1, ..., x_k$ совпадают, т.е. эти точки являются различными корнями разности $P_{k+1}(x) - P_k(x)$, то

$$P_{k+1}(x) - P_k(x) = A_{k+1} \cdot (x - x_0) \cdot (x - x_1) \cdot \ldots \cdot (x - x_k)\,, \tag{8}$$

где постоянная A_{k+1} определяется из уравнения

$$P_{k+1}(x_{k+1}) - P_k(x_{k+1}) = A_{k+1} \cdot (x_{k+1} - x_0) \cdot (x_{k+1} - x_1) \cdot \ldots \cdot (x_{k+1} - x_k)\,.$$

Так как $P_{k+1}(x_{k+1}) = f(x_{k+1})$, то

$$A_{k+1} = \frac{f(x_{k+1}) - P_k(x_{k+1})}{(x_{k+1} - x_0) \cdot (x_{k+1} - x_1) \cdot \ldots \cdot (x_{k+1} - x_k)}\,. \tag{9}$$

Легко получить явную формулу для A_1 :

$$A_1 = \frac{f(x_1) - f(x_0)}{x_1 - x_0} = \frac{f(x_0)}{x_0 - x_1} \;+\; \frac{f(x_1)}{x_1 - x_0}\,. \tag{10}$$

Определение 2. Разделенной разностью первого порядка функции $f(x)$ на узлах $x_0 \neq x_1$ называется число

$$f(x_0,\, x_1) = \frac{f(x_1) - f(x_0)}{x_1 - x_0}\,.$$

Тогда интерполяционный полином $P_1(x)$ можно представить в следующей форме (Ньютона):

$$\begin{aligned} P_1(x) &= P_0(x) \;+\; A_1 \cdot (x - x_0) = \\ &= f(x_0) \;+\; f(x_0,\, x_1) \cdot (x - x_0)\,. \end{aligned}$$

Определим интерполяционный полином $Q_1(x)$ для функции $f(x)$ на узлах $x_1 \neq x_2$ в форме (Ньютона):

$$Q_1(x) = f(x_1) \;+\; f(x_1,\, x_2) \cdot (x - x_1)\,.$$

Тогда интерполяционный полином $P_2(x)$ можно получить двумя способами:

$$\begin{aligned} P_2(x) = P_1(x) \ &+ \ (\,P_2(x) - P_1(x)\,) = \\ = P_1(x) \ &+ \ A_2 \cdot (x - x_0) \cdot (x - x_1) \end{aligned}$$

и

$$\begin{aligned} P_2(x) = Q_1(x) \ &+ \ (\,P_2(x) - Q_1(x)\,) = \\ = Q_1(x) \ &+ \ B_2 \cdot (x - x_1) \cdot (x - x_2)\,. \end{aligned}$$

Используя эти формулы при $x = x_2$, нетрудно найти выражение для A_2 :

$$\begin{aligned} A_2 &= \frac{Q_1(x_2) \ - \ P_1(x_2)}{(x_2 - x_0) \cdot (x_2 - x_1)} = \frac{f(x_1,\, x_2) \ - \ f(x_0,\, x_1)}{x_2 - x_0} = \\ &= \frac{f(x_0)}{(x_0 - x_1)\,(x_0 - x_2)} + \frac{f(x_1)}{(x_1 - x_0)\,(x_1 - x_2)} + \frac{f(x_2)}{(x_2 - x_0)\,(x_2 - x_1)}\,. \end{aligned}$$

Определение 3. Разделенной разностью второго порядка функции $f(x)$ на попарно различных узлах x_0, x_1, x_2 называется число

$$f(x_0,\, x_1,\, x_2) = \frac{f(x_1,\, x_2) \ - \ f(x_0,\, x_1)}{x_2 - x_0}\,.$$

Тогда интерполяционный полином $P_2(x)$ можно представить в следующей форме (Ньютона):

$$\begin{aligned} P_2(x) &= P_0(x) \ + \ A_1 \cdot (x - x_0) \ + \ A_2 \cdot (x - x_0) \cdot (x - x_1) = \\ &= f(x_0) + f(x_0,\, x_1)\,(x - x_0) + f(x_0,\, x_1,\, x_2)\,(x - x_0)\,(x - x_1)\,. \end{aligned}$$

Полученные формулы позволяют предположить, что коэффициенты A_k равны разделенным разностям k-того порядка функции $f(x)$ на попарно различных узлах $x_0, x_1, ..., x_k$, которые определяются через разности предыдущего порядка.

Определение 4. Разделенной разностью k-того порядка функции $f(x)$ на попарно различных узлах $x_0, x_1, ..., x_k$ называется число

$$f(x_0,\, x_1,\, \dots,\, x_k) = \frac{f(x_1,\, x_2 \dots,\, x_k) \ - \ f(x_0,\, x_1 \dots,\, x_{k-1})}{x_k - x_0}\,.$$

Лемма 1. *Для разделенной разностью k-того порядка функции $f(x)$ на попарно различных узлах $x_0, x_1, ..., x_k$ справедлива следующая формула:*

$$f(x_0,\, x_1,\, \dots,\, x_k) = \sum_{i=0}^{k} \frac{f(x_i)}{(x_i - x_0) \dots (x_i - x_{i-1})\,(x_i - x_{i+1}) \dots (x_i - x_k)}\,.$$

Доказательство. Справедливость леммы для $k = 1$ следует очевидным образом из определения разделенной разности первого порядка. Предлоложим, что лемма справедлива при $k = m$ и докажем ее для $k = m + 1$.

Из определения разделенной разности $(m + 1)$-го порядка и сделанного предположения математической индукции следует, что

$$f(x_0,\, x_1,\, \dots,\, x_{m+1}) = \frac{f(x_1,\, x_2 \dots,\, x_{m+1}) \,-\, f(x_0,\, x_1 \dots,\, x_m)}{x_{m+1} - x_0} =$$

$$= \frac{1}{x_{m+1} - x_0} \Big[\sum_{i=1}^{m+1} \frac{f(x_i)}{(x_i - x_1) \dots (x_i - x_{i-1})\,(x_i - x_{i+1}) \dots (x_i - x_{m+1})} -$$

$$- \sum_{i=0}^{m} \frac{f(x_i)}{(x_i - x_0) \dots (x_i - x_{i-1})\,(x_i - x_{i+1}) \dots (x_i - x_m)} \Big] =$$

$$= \sum_{i=0}^{m+1} \frac{f(x_i)}{(x_i - x_0) \dots (x_i - x_{i-1})\,(x_i - x_{i+1}) \dots (x_i - x_{m+1})}\ ,$$

что и требовалось доказать. □

Лемма 2. *Коэффициенты A_k из (9) равны разделенным разностям k-того порядка функции $f(x)$ на попарно различных узлах $x_0, x_1, \dots, x_k$.*

Доказательство. Так как легко установить справедливось формулы

$$A_k = \sum_{i=0}^{k} \frac{f(x_i)}{(x_i - x_0) \dots (x_i - x_{i-1})\,(x_i - x_{i+1}) \dots (x_i - x_k)}$$

(заменой в формуле (9) интерполяционного полинома на его представление в форме Лагранжа и приведением подобных), то утверждение леммы будет следствием леммы 1. □

Теорема 3. *Решение задачи алгебраической интерполяции (2) при $m = n$ представимо в форме Ньютона:*

$$P_n(x) = \sum_{k=0}^{n} f(x_0,\, x_1,\, \dots,\, x_k) \cdot (x - x_0) \cdot (x - x_1) \cdot \ldots \cdot (x - x_{k-1})\,,$$

где $f(x_0,\, x_1,\, \dots,\, x_k)$ – разделенные разности k-того порядка.

Доказательство. Утверждение теоремы является следствием формулы (8) и леммы 2. □

1.3. Оценка погрешности интерполирования

Определение 5. Погрешностью (ошибкой) интерполирования функции $f(x)$ по ее значениям $f(x_i)$ в попарно различных точках $x_0, x_1, \dots, x_n$ интервала $[\,a, b\,]$ интерполяционным полиномом $P_n(x)$ называется их разность

$$R_n(x) = f(x) - P_n(x)\,.$$

Очевидно, что $R_n(x_i) = 0$.

Лемма 3. *Если $y \in [\,a,\ \ b\,]$ и не совпадает ни с одним из узлов интерполяции, то*

$$R_n(y) \;=\; f(x_0,\, \ldots\,,\, x_n,\, y) \cdot \omega(y)\,,$$

где $\omega(y) = (y - x_0) \cdot (y - x_1) \cdot \ldots \cdot (y - x_n)$.

Доказательство. Интерполяционный полином $P_{n+1}(x)$ функции $f(x)$ по попарно различным узлам интерполяции $x_0, x_1, \ldots, x_n, y$ представим в форме Ньютона:

$$P_{n+1}(x) = f(x_0) + \ldots + f(x_0,\, x_1,\, \ldots\,,\, x_n)\,(x - x_0)\,(x - x_1)\ldots(x - x_{n-1}) + \\ + f(x_0,\, \ldots\,,\, x_n,\, y)\,\omega(x) = P_n(x) + f(x_0,\, \ldots\,,\, x_n,\, y)\,\omega(x)\,.$$

Утверждение леммы следует из этого равенства и условия интерполяции $P_{n+1}(y) = f(y)$. □

Лемма 4. *Если функция $f(x)$ n раз непрерывно дифференцируема на отрезке $[\,a,\ \ b\,]$, то*

$$f(x_0,\, x_1,\, \ldots\,,\, x_n) \;=\; \frac{f^{(n)}(\xi)}{n!}$$

для некоторой точки $\xi \in [\,a,\ \ b\,]$.

Доказательство. Прежде всего продифференцируем n раз ошибку интерполяции:

$$R_n^{(n)}(x) \;=\; f^{(n)}(x) \;-\; n! \cdot f(x_0,\, x_1,\, \ldots\,,\, x_n)\,.$$

Очевидно, что для доказательства леммы достаточно установить существование хотя бы одного корня ξ у этой производной на $[\,a,\ \ b\,]$.

Подсчитаем количество различных корней у погрешности $R_n(x)$: их не меньше, чем $n+1$, так как $R_n(x_i) = 0\,,\ \ i = 0, 1, \ldots, n$. Не уменьшая общности, будем считать, что узлы интерполяции упорядочены по возрастанию: $x_0 < x_1 < \ldots < x_n$. Тогда по теореме о среднем значении ее первая производная будет иметь хотя бы один корень внутри каждого интервала $(\,x_0,\, x_1\,)$, $(\,x_1,\, x_2\,)$, ..., $(\,x_{n-1},\, x_n\,)$. Следовательно, $R_n'(x)$ имеет на $[\,a,\ \ b\,]$ не менее, чем n попарно различных корней.

Повторяя это рассуждение для функции $R_n'(x)$, получим, что вторая производная погрешности $R_n''(x)$ имеет на $[\,a,\ \ b\,]$ не менее, чем $n-1$ попарно различных корней. Через n таких шагов мы придем к выводу, что n-тая производная погрешности $R_n^{(n)}(x)$ имеет на $[\,a,\ \ b\,]$ не менее одного корня, что и требовалось доказать. □

Теорема 4. *Если функция $f(x)$ $n+1$ раз непрерывно дифференцируема на отрезке $[\,a,\ \ b\,]$, то погрешность интерполирования функции $f(x)$ по ее значениям $f(x_i)$ в попарно различных точках $x_0, x_1, \ldots, x_n$ интервала $[\,a, b\,]$, интерполяционным полиномом $P_n(x)$ может быть представлена в следующем виде:*

$$R_n(y) = f(y) - P_n(y) = \frac{f^{(n+1)}(\xi)}{(n+1)!}\,\omega(y)\,,$$

где $\xi \in [\,a,\ \ b\,]$, $\omega(y) = (y - x_0) \cdot (y - x_1) \cdot \ldots \cdot (y - x_n)$.

Доказательство. Если y совпадает с одним из узлов интерполирования, то, утверждение леммы очевидным образом верно для любого $\xi \in [\,a, b\,]$, поскольку $R_n() = 0$ и $\omega(y) = 0$.

Если y не совпадает ни с одним из узлов интерполирования, то из леммы 3 следует, что
$$R_n(y) \;=\; f(x_0,\, \ldots\,,\, x_n,\, y) \cdot \omega(y)\,,$$
а лемма 4 гарантирует существование точки $\xi \in [\,a, b\,]$, такой, что
$$f(x_0,\, x_1,\, \ldots\,,\, x_n,\, y) \;=\; \frac{f^{(n+1)}(\xi)}{(n+1)!}\,,$$
следовательно,
$$R_n(y) \;=\; \frac{f^{(n+1)}(\xi)}{(n+1)!}\,\omega(y)$$
и в этом случае. $\square$

2. Интерполирование с кратными узлами

Иногда требуется построить полиномиальное приближение функции $f(x)$ при условии, что известны не только ее значения в некоторых точках области задания, но и значения ее производных.

Определение 1. Алгебраический полином

$$P_N(x) = a_0 + a_1 \cdot x + ... + a_N \cdot x^N \tag{1}$$

называется интерполяционным полиномом Эрмита для функции $f(x)$, заданной на отрезке $[\,a,\ b\,]$, по ее значениям $f(x_i)$, значениям ее первых $N_i - 1$ производных $f^{(j)}(x_i)$ в $(n+1)$-ой попарно различных точках (узлах) $x_i \in [\,a,\ b\,]$, $i = 0, 1, ..., n$, если

$$\begin{gathered} P_N^{(j)}(x_i) = f^{(j)}(x_i), \\ i = 0, 1, ..., n\,; \quad j = 0, 1, ..., N_i - 1\,. \end{gathered} \tag{2}$$

Точки x_i называются узлами интерполяции кратности N_i. □

Для определения полинома $P_N(x)$ ($N + 1$ неизвестное: $a_0,\ a_1, ...,\ a_N$) мы имеем $N_0 + N_1 + ... + N_n$ условий (уравнения (2)). Единственность решения математических конечномерных задач обычно обеспечивается равенством числа неизвестных количеству накладываемых на них условий, поэтому впредь будем считать, что $N + 1 = N_0 + N_1 + ... + N_n$.

Теорема 1. *Задача алгебраической интерполяции (2) при $N+1 = N_0+N_1+ ... + N_n$ имеет единственное решение.*

Доказательство. Очевидно, что система уравнений (2) относительно неизвестных $a_0,\ a_1, ...,\ a_N$ является линейной алгебраической системой порядка $N + 1$. Следовательно она имеет единственное решение в том и только том случае, когда однородная система имеет только нулевое решение (определитель матрицы этой системы отличен от нуля).

Предположим, что теорема неверна: система уравнений (2) либо не имеет решения, либо имеет несколько решений. Значит однородная система имеет ненулевое решение $a_0,\ a_1, ..., a_N$, определяющее ненулевой полином $P_N(x)$ степени N и удовлетворяющий условиям

$$P_N^{(j)}(x_i) = 0, \quad i = 0, 1, ..., n\,; \quad j = 0, 1, ..., N_i - 1\,.$$

Отсюда следует, что $x = x_i$ является его корнем кратности N_i, а, так как интерполяционные узлы $x = x_i$ попарно различны, то

$$P_N(x) = c\,(x - x_0)^{N_0}\,(x - x_1)^{N_1} \dots (x - x_n)^{N_n}\,,$$

т.е. $P_N(x)$ является полиномом степени не менее, чем $N_0 + N_1 + ... + N_n$, так как по построению он не является тождественно нулевым. Но по условия теоремы его степень строго меньше $N_0 + N_1 + ... + N_n$, следовательно, наше предположение ложно и теорема верна. □

2.1. Представления полинома Эрмита

Форму Лагранжа представления интерполяционного полинома для случая простых интерполяционных узлов можно обощить и на случай кратных узлов в следующем виде

$$P_N(x) = \sum_{i=0}^{n} \sum_{j=0}^{N_i-1} P_{N,i,j}(x) \cdot f^{(j)}(x_i)\,, \tag{3}$$

где полином $P_{N,i,j}(x)$ степени N определяется условиями

$$P^{(l)}_{N,i,j}(x_k) = \delta_{i,k} \cdot \delta_{j,l}\,,$$
$$k = 0, 1, ..., n\,; \quad l = 0, 1, ..., N_i - 1\,,$$

где $\delta_{i,k}$ – символ Кронекера.

Пример. Построим полином Эрмита при $n = 1$, $N_0 = 1$, $N_1 = 2$. Введем дополнительный узел x_2, отличный от x_0 и x_1, и построим интерполяционный полином $L_2(x)$ в форме Лагранжа на узлах x_0, x_1, x_2

$$L_2(x) = \frac{(x-x_1)\,(x-x_2)}{(x_0-x_1)\,(x_0-x_2)} f(x_0) + \frac{(x-x_0)\,(x-x_2)}{(x_1-x_0)\,(x_1-x_2)} f(x_1) + \\ + \frac{(x-x_0)\,(x-x_1)}{(x_2-x_0)\,(x_2-x_1)} f(x_2)\,.$$

Предельным переходом при $x_2 \to x_1$ получим полином Эрмита

$$P_2(x) = \frac{(x-x_1)^2}{(x_0-x_1)^2} f(x_0) - \frac{(x-x_0)\,(x+x_0-2x_2)}{(x_1-x_0)^2} f(x_1) + \\ + \frac{(x-x_0)\,(x-x_1)}{x_1-x_0} f'(x_1)\,.$$

Интерполяционный полином Эрмита можно построить и с помощью предельного перехода в интерполяционном полиноме в форме Ньютона. Построим интерполяционный полином $L_2(x)$ в форме Ньютона на узлах x_0, x_1, x_2

$$L_2(x) = f(x_0) + f(x_0,\, x_1)\,(x-x_0) + f(x_0,\, x_1,\, x_2)\,(x-x_0)\,(x-x_1)\,,$$
$$f(x_0,\, x_1) = \frac{f(x_1) - f(x_0)}{(x_1-x_0)}\,, \quad f(x_1,\, x_2) = \frac{f(x_2) - f(x_1)}{(x_2-x_1)}\,,$$
$$f(x_0,\, x_1,\, x_2) = \frac{f(x_1,\, x_2) - f(x_0,\, x_1)}{(x_2-x_0)}\,.$$

Предельным переходом при $x_2 \to x_1$ получим полином Эрмита

$$P_2(x) = f(x_0) + f(x_0,\, x_1)\,(x-x_0) + f(x_0,\, x_1,\, x_1)\,(x-x_0)\,(x-x_1) = \\ = f(x_0) + \frac{f(x_1) - f(x_0)}{(x_1-x_0)}\,(x-x_0) + \\ + \frac{(x_1-x_0)\,f'(x_1) - f(x_1) + f(x_0)}{(x_1-x_0)^2}\,(x-x_0)\,(x-x_1)\,.$$

Легко заметить, что в этом случае находить предел значительно проще, чем при использовании полинома в форме Лагранжа. □

Установим законность последнего приема для построения интерполяционного полинома Эрмита.

Предположим, что каждый узел интерполяции x_i кратности N_i заменен набором $x^\varepsilon_{i,0},\ x^\varepsilon_{i,1},\ \dots,\ x^\varepsilon_{i,N_i-1}$, зависящим от параметра ε, при стремлении которого к нулю каждый узел набора сходиться к x_i.

Будем считать, что узлы набора

$$x^\varepsilon_{i,0},\ x^\varepsilon_{i,1},\ \dots,\ x^\varepsilon_{i,N_i-1},\ i=0,1,\dots,n,$$

попарно различны (построить такую систему узлов можно, выбирая дополнительные узлы в достаточно малых ε-окрестностях исходных интерполяционных узлов).

Представим интерполяционный полином на этой системе узлов в форме Ньютона

$$\begin{aligned}
P^\varepsilon_N(x) = A^\varepsilon_{0,0}\ &+ \\
+\ &A^\varepsilon_{0,1}(x-x^\varepsilon_{0,0})\ + \\
+\ &A^\varepsilon_{0,2}(x-x^\varepsilon_{0,0})(x-x^\varepsilon_{0,1})\ + \\
&\dots\dots\dots\dots \\
+\ &A^\varepsilon_{0,N_0-1}(x-x^\varepsilon_{0,0})\dots(x-x^\varepsilon_{0,N_0-2})\ + \\
+\ &A^\varepsilon_{1,0}(x-x^\varepsilon_{0,0})\dots(x-x^\varepsilon_{0,N_0-2})(x-x^\varepsilon_{0,N_0-1})\ + \\
&\dots\dots\dots\dots \\
+\ &A^\varepsilon_{n,N_n-1}(x-x^\varepsilon_{0,0})\dots(x-x^\varepsilon_{0,N_0-2})(x-x^\varepsilon_{0,N_0-1})\dots(x-x^\varepsilon_{n,N_n-2}),
\end{aligned}$$

где

$$\begin{aligned}
A^\varepsilon_{0,0} &= f(x^\varepsilon_{0,0}), \\
A^\varepsilon_{0,1} &= f(x^\varepsilon_{0,0},\, x^\varepsilon_{0,1}), \\
&\dots\dots\dots \\
A^\varepsilon_{0,N_0-1} &= f(x^\varepsilon_{0,0},\, x^\varepsilon_{0,1},\, \dots,\, x^\varepsilon_{0,N_0-1}), \\
A^\varepsilon_{1,0} &= f(x^\varepsilon_{0,0},\, x^\varepsilon_{0,1},\, \dots,\, x^\varepsilon_{0,N_0-1},\, x^\varepsilon_{1,0}), \\
&\dots\dots\dots\dots \\
A^\varepsilon_{n,N_n-1} &= f(x^\varepsilon_{0,0},\, x^\varepsilon_{0,1},\, \dots,\, x^\varepsilon_{0,N_0-1},\, x^\varepsilon_{1,0},\, \dots,\, x^\varepsilon_{n,N_n-1}).
\end{aligned}$$

Лемма 1. *На системе узлов* $\{x^\varepsilon_{i,0},\ x^\varepsilon_{i,1},\ \dots,\ x^\varepsilon_{i,N_i-1},\ i=0,1,\dots,n\}$ *существует и конечен предел разделенных разностей:*

$$\lim_{\varepsilon\to 0} f(x^\varepsilon_{i,j},\, x^\varepsilon_{i,j+1},\, \dots,\, x^\varepsilon_{k,l}) \equiv f(x_{i,j},\, x_{i,j+1},\, \dots,\, x_{k,l}),$$

$$0\le i\le k\le n,\ \ 0\le j\le N_i-1,\ \ 0\le l\le N_k-1,$$

где $\lim\limits_{\varepsilon\to 0} x^\varepsilon_{i,j} = x_{i,j} \equiv x_i$.

Доказательство. Очевидно, что пределы разделенных разностей нулевого порядка существуют и конечны:

$$\lim_{\varepsilon\to 0} f(x^\varepsilon_{i,j}) = f(x_{i,j}) \equiv f(x_i)$$

$$0\le i\le k\le n,\ \ 0\le j\le N_i-1,$$

Предположим, что существуют и конечны пределы разделенных разностей до порядка m включительно и рассмотрим разделенную разность порядка $m+1$:

$$f(x^{\varepsilon}_{i,j},\, x^{\varepsilon}_{i,j+1},\, ...,\, x^{\varepsilon}_{k,l}) \;=\; \frac{f(x^{\varepsilon}_{i,j+1},\, ...,\, x^{\varepsilon}_{k,l}) \;-\; f(x^{\varepsilon}_{i,j},\, x^{\varepsilon}_{i,j+1},\, ...,\, x^{\varepsilon}_{k,l-1})}{x^{\varepsilon}_{k,l} \;-\; x^{\varepsilon}_{i,j}}\,,$$

где числитель является разностью разделенных разностей m-го порядка, а значит по предположению имеет конечный предел.

Если $i \;\neq\; k$, то знаменатель стремиться к разности $x_k - x_i \;\neq\; 0$ и нужный предел существует. Осталось рассмотреть случай $i \;=\; k$, когда знаменатель стремиться к нулю. Но в этом случае из леммы (1.4) следует, что

$$f(x^{\varepsilon}_{i,j},\, x^{\varepsilon}_{i,j+1},\, ...,\, x^{\varepsilon}_{i,l}) \;=\; \frac{f^{(m+1)}(\xi^{\varepsilon})}{(m+1)!}$$

и ξ^{ε} лежит между узлами, по которым строиться разделенная разность. Значит $\xi^{\varepsilon} \to x_i$, а разделенная разность стремиться к $(m+1)$-ой производной функции $f(x)$ в точке x_i , что и требовалось доказать. □

Теорема 2. *Интерполяционный полином $P^{\varepsilon}_N(x)$ на системе узлов*

$$\{\, x^{\varepsilon}_{i,0},\;\; x^{\varepsilon}_{i,1},\;\; \dots,\;\; x^{\varepsilon}_{i,N_i-1}\,,\;\; i = 0, 1, ..., n\,\}$$

сходится к полиному Эрмита $P_N(x)$ из (2) при $\varepsilon \to 0$.

Доказательство. Из леммы (1) и ее доказательства следует, что предел интерполяционного полинома $P^{\varepsilon}_N(x)$ можно представить в следующем виде:

$$\begin{aligned} P_N(x) = f(x_0) \,+\, \frac{f\,'(x_0)}{1!}(x - x_0) \,+\, ... \,+\, \frac{f^{(N_0-1)}(x_0)}{(N_0-1)!}(x - x_0)^{N_0-1} \,+ \\ +\, f(x_{0,0},\, \dots,\, x_{0,N_0-1},\, x_{1,0})\,(x - x_0)^{N_0} \,+ \\ +\, f(x_{0,0},\, \dots,\, x_{0,N_0-1},\, x_{1,0},\, x_{1,1})\,(x - x_0)^{N_0}\,(x - x_1) \,+ \\ .\;.\;.\;.\;.\;.\;.\;.\;.\;. \\ +\, f(x_{0,0},\, \dots,\, x_{n,N_n-1})\,(x - x_0)^{N_0}(x - x_1)^{N_1} ... \,(x - x_n)^{N_n-1}\,. \end{aligned}$$

Непосредственной проверкой легко установить, что этот предельный полином удовлетворяет всем условиям интерполяции функции $f(x)$ в интерполяционном узле x_0 кратности N_0 :

$$P_N(x_0) = f(x_0)\,,\;\; P'_N(x_0) = f\,'(x_0)\,,\;\; \dots\,,\;\; P_N^{(N_0-1)}(x_0) = f^{(N_0-1)}(x_0)\,.$$

Предельный полином $P_N(x)$ удовлетворяет условиям интерполяции функции $f(x)$ и в любом другом интерполяционном узле x_i кратности N_i . Действительно, поменяв узлы x_i и x_0 местами, учитывая, что интерполяционный полином $P^{\varepsilon}_N(x)$ единственен, т.е. не зависит от нумерации узлов, получим, что его предел представим в виде

$$\begin{aligned} P_N(x) = f(x_i) \,+\, &\frac{f'(x_i)}{1!}(x-x_i) \,+\, \dots \,+\, \frac{f^{(N_i-1)}(x_i)}{(N_i-1)!}(x-x_i)^{N_i-1} + \\ &+ f(x_{i,0},\, \dots,\, x_{i,N_i-1},\, x_{1,0})\,(x-x_i)^{N_i} + \\ &+ f(x_{i,0},\, \dots,\, x_{i,N_i-1},\, x_{1,0},\, x_{1,1})\,(x-x_i)^{N_i}\,(x-x_1) + \\ &\qquad \cdots\cdots\cdots \\ &+ f(x_{i,0},\, \dots,\, x_{n,N_n-1})\,(x-x_i)^{N_i}(x-x_1)^{N_1}\dots(x-x_n)^{N_n-1} \end{aligned}$$

и удовлетворяет условиям интерполяции

$$P_N(x_i) = f(x_i)\,,\;\; P_N'(x_i) = f'(x_i)\,,\;\; \dots,\;\; P_N^{(N_i-1)}(x_i) = f^{(N_i-1)}(x_i)\,.$$

2.2. Оценка погрешности интерполирования

Погрешность интерполирования функции $f(x)$, заданной на отрезке $[\,a,\;b\,]$, по ее значениям $f(x_i)$, значениям ее первых N_i-1 производных $f^{(j)}(x_i)$ в $(n+1)$-ой попарно различных точках (узлах) $x_i \in [\,a,\;b\,]$, $i = 0,1,\dots,n$, полиномом Эрмита $P_N(x)$ из (2)

$$R_N(x) = f(x) - P_N(x) \tag{4}$$

можно оценить с помощью предельного перехода в оценке

$$R_N^\varepsilon(x) = f(x) - P_N^\varepsilon(x) = \frac{f^{(N+1)}(\xi)}{(N+1)!)}\omega^\varepsilon(x)\,, \tag{5}$$

где $P_N^\varepsilon(x)$ – интерполяционный многочлен на попарно различных узлах кратности один,

$$\omega^\varepsilon(x) = (x-x_{0,0}^\varepsilon)(x-x_{0,1}^\varepsilon)\dots(x-x_{0,N_0-1}^\varepsilon)(x-x_{1,0}^\varepsilon)\dots(x-x_{n,N_n-1}^\varepsilon)\,.$$

Так как $P_N^\varepsilon(x) \to P_N(x)$, $\omega^\varepsilon(x) \to \omega(x)$ при $\varepsilon \to 0$:

$$\omega(x) = (x-x_0)^{N_0}(x-x_1)^{N_1}\dots(x-x_n)^{N_n} \tag{6}$$

– полином степени $N+1 = N_0+N_1+\dots+N_n$, то из представления (5) следует

$$||R_N(x)||_{C[a,b]} \;\le\; \frac{||f^{(N+1)}(x)||_{C[a,b]}}{(N+1)!}\,||\omega(x)||_{C[a,b]}\,. \tag{7}$$

Эта оценка вытекает также и из следующей теоремы.

Теорема 3. *Погрешность $R_N(x)$ интерполирования функции $f(x)$, заданной на отрезке $[\,a,\;b\,]$, по ее значениям $f(x_i)$, значениям ее первых N_i-1 производных $f^{(j)}(x_i)$ в $(n+1)$-ой попарно различных точках (узлах) $x_i \in [\,a,\;b\,]$, $i = 0,1,\dots,n$, полиномом Эрмита $P_N(x)$ из (2) можно представить в виде*

$$R_N(x) = f(x) - P_N(x) = \frac{f^{(N+1)}(\xi)}{(N+1)!)}\omega(x)\,,$$

где ξ некоторая точка из интервала $[\,a,\;b\,]$ (зависящая от x), а полином $\omega(x)$ определен в (6).

Доказательство. Очевидно, что, если $x = y$ и y совпадает с одним из узлов интерполяции, то утверждение теоремы справедливо при любом выборе $\xi \in [\,a,\ b\,]$, так как $R_N(y) = 0$ и $\omega(y) = 0$.

Пусть $x = y$ и y не совпадает ни с одним из узлов интерполяции. Определим постоянную

$$K \;=\; \frac{f(y) - P_N(y)}{\omega(y)}\,,$$

тогда функция $g(x) = R_N(x) - K \cdot P_N(x)$ имеет по крайней мере $N+1$ корень (вместе с их кратностями) $x_0,\, x_1,\, \dots,\, x_n$ по условиям интерполяции и еще один корень $x = y$ по определению постоянной K, всего $N + 2$ корня.

Очевидно, что первая производная функции $g(x)$ на интервале $[\,a,\ b\,]$ будет иметь корни $x_0,\, x_1,\, \dots,\, x_n$ кратности $N_0 - 1,\, N_1 - 1,\, \dots,\, N_n - 1$ соответственно, кроме того, по теореме Ролля еще $n + 1$ корень внутри подъинтервалов разбиения отрезка $[\,a,\ b\,]$ точками $x_0,\, x_1,\, \dots,\, x_n$ и y, т.е. $N + 1$ корень.

Повторяя это рассуждение необходимое число раз, приходим к выводу, что l-ая производная функции $g(x)$ будет иметь на интервале $[\,a,\ b\,]$ по крайней мере $N + 2 - l$ корней и, следовательно, существует точка $\xi \in [\,a,\ b\,]$ такая, что

$$0 \;=\; g^{(N+1)}(\xi) \;=\; R_N^{(N+1)}(\xi) \;-\; K \cdot (N+1)! \;=\; f^{(N+1)}(\xi) \;-\; K \cdot (N+1)!\,,$$

т.е.

$$K \;=\; \frac{f(y) - P_N(y)}{\omega(y)} \;=\; f^{(N+1)}(\xi)\,,$$

откуда следует утверждение теоремы. □

3. Полином (Чебышева), наименее уклоняющийся от нуля

Рассмотрим оценку (2.7) ошибки интерполяции функции $f(x)$

$$||R_N(x)||_{C[a,b]} \le \frac{||f^{(N+1)}(x)||_{C[a,b]}}{(N+1)!} ||\omega(x)||_{C[a,b]} ,$$

где

$$\omega(x) = (x - x_0)^{N_0}(x - x_1)^{N_1} \dots (x - x_n)^{N_n}$$

– полином степени $N + 1 = N_0 + N_1 + \dots + N_n$,
$x_0, x_1, \dots , x_n$ – интерполяционные узлы кратности $N_0, N_1, \dots , N_n$ соответственно.

Неудачный выбор узлов интерполяции (например, в одной половине интервала $[\,a,\ b\,]$) будет причиной плохого приближения функции $f(x)$ интерполяционным полиномом (в другой половине отрезка $[\,a,\ b\,]$). Поэтому естественной является задача такого выбора интерполяционных узлов, при котором норма многочлена $\omega(x)$ в пространстве $C\,[\,a,\ b\,]$ минимальна.

Мы ограничимся случаем простых ($N_i = 1$) узлов интерполяции, т.е.

$$\begin{aligned} \omega(x) &= (x - x_0)\,(x - x_1) \dots (x - x_n) = \\ &= x^{n+1} - (p_n x^n + \dots + p_1 x + p_0) = x^{n+1} - P_n(x) \end{aligned} \tag{1}$$

– полином степени $n + 1$ с равным единице коэффициентом при старшей степени x^{n+1} равным единице.

Легко понять, что речь идет о наилучшем приближении функции $f(x) = x^{n+1}$ полиномом $P_n(x)$ степени n на отрезке $[\,a,\ b\,]$:

$$||x^{n+1} - P_n(x)||_{C[a,b]} = \inf_{Q_n(x)} ||x^{n+1} - Q_n(x)||_{C[a,b]} , \tag{2}$$

а корни многочлена $\omega(x)$ должны быть попарно различны и принадлежать интервалу $[\,a,\ b\,]$.

3.1. Наилучшие приближения в нормированных пространствах

Теорема 1. *Пусть B – нормированное пространство, B_{n+1} – его подпространство размерности $n+1$. Тогда для любого элемента $f \in B$ существует $P_n \in B_{n+1}$ такой, что*

$$||f - P_n||_B = \inf_{Q_n \in B_{n+1}} ||f - Q_n||_B . \tag{3}$$

Доказательство. Пусть $\{b_0, b_1, \dots , b_n\}$ – базис B_{n+1} , т.е. любой элемент $Q_n \in B_{n+1}$ может быть представлен в виде линейной комбинации

$$Q_n = q_0 \cdot b_0 + q_1 \cdot b_1 + \ldots + q_n \cdot b_n$$

с вектором коэффициентов $\vec{q} = (q_0, q_1, \ldots, q_n)$ из векторного пространства R^{n+1}.

Определим в R^{n+1} неотрицательную функцию

$$R_f(\vec{q}) = ||\, f - Q_n \,||_B \,.$$

Очевидно, что задача (3) эквивалентна поиску точки минимума этой функции в R^{n+1} :

$$R_f(\vec{p}) = \inf_{\vec{q} \in R^{n+1}} R_f(\vec{q}) = \alpha \geq 0 \,.$$

Известно, что, если функция непрерывна и ее минимум ищется на компактном множестве, то он существует и достигается.

Непрерывность функции $R_f(\vec{q})$ очевидна.

Компактными множествами в R^{n+1} являются замкнутые шары. Покажем, что минимум $R_f(\vec{q})$ не может достигаться вне некоторого шара $||\,\vec{q}\,||_{R^{n+1}} \leq C_f$.

Действительно, так как

$$||\, Q_n \,||_B = ||\,\vec{q}\,||_{R^{n+1}} \cdot ||\sum_{i=0}^{n} \beta_i \cdot b_i \,||_B \,,$$
$$||\,\vec{\beta}\,||_{R^{n+1}} = 1 \,,$$
$$\inf_{||\vec{\beta}||_{R^{n+1}}=1} ||\sum_{i=0}^{n} \beta_i \cdot b_i \,||_B = \beta > 0 \,,$$
$$R_f(\vec{q}) \geq ||\,\vec{q}\,||_{R^{n+1}} \cdot \beta - ||\, f \,||_B \,,$$

то в качестве C_f можно взять $(\alpha + ||f||_B)/\beta$. $\square$

Итак, наилучшее приближение элемента f нормированного пространства B элементом из его конечномерного подпространства существует, но единственность такого приближения теорема 1 не гарантирует.

Теорема 2. *Множество решений задачи о наилучшем приближении (3) выпукло.*

Доказательство. Пусть $P_n^{(1)} \in B_{n+1}$ и $P_n^{(2)} \in B_{n+1}$ – любые два решения задачи (3), т.е.

$$||f - P_n^{(1)}||_B = ||f - P_n^{(2)}||_B = \inf_{Q_n \in B_{n+1}} ||f - Q_n||_B = \alpha \,.$$

Тогда для $Q_n = (1-t) \cdot P_n^{(1)} + t \cdot P_n^{(1)}$ при любом $t \in [0, 1]$ имеем

$$\begin{aligned} \alpha \le ||f - Q_n||_B &= ||(1-t)(f - P_n^{(1)}) + t(f - P_n^{(2)})||_B \le \\ &\le (1-t)\cdot ||f - P_n^{(1)}||_B + t\cdot ||f - P_n^{(2)}||_B = \\ &= (1-t)\cdot \alpha + t\cdot \alpha = \alpha\,, \end{aligned} \tag{4}$$

т.е. любая точка отрезка, соединяющего точки $P_n^{(1)}$ и $P_n^{(2)}$ в B_{n+1}, является наилучшим приближением элемента f, что означает выпуклость множества решений задачи (3). □

Определение 1. Нормированное пространство B называется строго нормированным, если для любых линейно независимых его элементов f и g выполняется строгое неравенство

$$||(1-t)\cdot f + t\cdot g||_B < (1-t)\cdot ||f||_B + t\cdot ||g||_B$$

при любом $t \in (0,\, 1)$. □

Теорема 3. *В строго нормированном пространстве B решение задачи о наилучшем приближении (3) единственно.*

Доказательство. Предположим, что решений задачи (3) два: $P_n^{(1)} \ne P_n^{(2)}$. Рассмотрим разности $f - P_n^{(1)}$ и $f - P_n^{(2)}$, норма которых одинакова и равна α. Если эти разности линейно независимы, то при любом $t \in (0,\, 1)$ из теоремы 1 и определения 1 имеем

$$\begin{aligned} \alpha &= ||f - [(1-t)P_n^{(1)}) + tP_n^{(2)}]||_B = \\ &= ||(1-t)(f - P_n^{(1)}) + t(f - P_n^{(2)})||_B < \\ &< (1-t)\cdot ||f - P_n^{(1)}||_B + t\cdot ||f - P_n^{(2)}||_B = \\ &= (1-t)\cdot \alpha + t\cdot \alpha = \alpha\,, \end{aligned}$$

Полученное противоречие $\alpha < \alpha$ доказывает единственность наилучшего приближения в этом случае.

Осталось рассмотреть случай $f - P_n^{(2)} = a\cdot (f - P_n^{(1)})$, где $a \ne 1$. Из теоремы 1 при любом $t \in (0,\, 1)$ следует

$$\begin{aligned} \alpha &= ||\, f - [(1-t)\cdot P_n^{(1)}) + t\cdot P_n^{(2)}]\, ||_B = \\ &= ||\,(1-t)\cdot (f - P_n^{(1)}) + t\cdot (f - P_n^{(2)})\, ||_B = \\ &= ||\,[(1-t) + t\cdot a]\cdot (f - P_n^{(1)})\, ||_B = \\ &= |\,(1-t) + t\cdot a\,|\cdot \alpha\,. \end{aligned}$$

Так как $a \ne 1$, то это равенство возможно лишь при $\alpha = 0$, т.е. $P_n^{(1)} = P_n^{(2)}$. □

Теорема 4. *В гильбертовом пространстве решение задачи о наилучшем приближении (3) существует и единственно.*

Доказательство. Существование наилучшего приближения следует из теоремы 1, так как гильбертово пространство является нормированным.

В соответствии с теоремой 3 для доказательства его единственности достаточно проверить строгую нормированность гильбертова пространства.

Так как норма $||f||_B$ в гильбертовом пространстве B определяется определяется через скалярное произведение $(f,\, g)_B$ формулой

$$|| \, f \, ||_B^2 \, = \, (\, f, \, g \,)_B \, ,$$

то

$$\begin{aligned} &||(1-t)\cdot f \, + t \cdot g||_B^2 \, = \\ &\qquad = (1-t)^2 \cdot ||f||_B^2 \, + \, (1-t)\cdot t \cdot ||f||_B \cdot ||g||_B \, + \, t^2 \cdot ||g||_B^2 \, < \\ &\qquad < [\,(1-t)\cdot ||f||_B \, + \, t \cdot ||g||_B \,]^2 \, , \end{aligned}$$

поскольку $(f,\, g)_B \, < \, ||f||_B \cdot ||g||_B$ для любых линейно независимых элементов f и g, и, следовательно, гильбертово пространство строго нормировано. □

К сожалению применить теорему 3 для доказательства единственности полинома наилучшего приближения в равномерной норме нельзя, так как пространство $C[a,\, b]$ не является строго нормированным: легко проверить, что для функций $f(x) = x$ и $g(x) = x^2$ из $C[0,\, 1]$ неравенство из определения 1 не выполняется.

3.2. Чебышевский альтернанс

Вернемся к задаче (2) поиска наилучшего приближения функции $f(x) = x^{n+1} \in C[\,a,\; b\,]$ полиномом $P_n(x)$ степени n :

$$||x^{n+1} \, - \, P_n(x)||_{C[a,b]} \, = \, \min_{Q_n(x) \in B_{n+1}} ||x^{n+1} \, - \, Q_n(x)||_{C[a,b]} \, , \tag{5}$$

где B_{n+1} – линейная оболочка функций $\{\,1,\, x,\, \dots,\, x^n\,\}$.

Найдем ее решение $P_0(x) = p_0$ при $n = 0$, когда $a = -1,\; b = 1$.
Так как $||x - q_0||_{C[-1,\, 1]} = \max\{|1+q_0|, |1-q_0|\} \geq 1$, то, очевидно, что $P_0(x) = 0$. Заметим, что максимальное по модулю значение L разности $x - P_0(x)$ достигается только в двух точках $y_0 = -1 < y_1 = 1$:

$$\begin{aligned} (x - P_0(x))|_{x=y_0} &= -L \, , \\ (x - P_0(x))|_{x=y_1} &= +L \, . \end{aligned}$$

Можно проверить, что решением задачи (5) при $n = 1$ будет $P_1(x) = 0.5$, а максимальное по модулю значение L разности $x^2 - P_1(x)$ достигается только в трех точках $y_0 = -1 < y_1 = 0 < y_2 = 1$:

$$\begin{aligned} (x - P_1(x))|_{x=y_0} &= +L \, , \\ (x - P_1(x))|_{x=y_1} &= -L \, , \\ (x - P_1(x))|_{x=y_2} &= +L \, . \end{aligned}$$

Определение 2. Говорят, что разность $f(x) - P_n(x)$ имеет чебышевский альтернанс в точках $y_0 < y_1 < \ldots < y_{n+1}$ из интервала $[\,a,\ b\,]$, если в этих точках она принимает максимальное по модулю значение, а знаки значений $f(y_0),\ f(y_1),\ \ldots,\ f(y_{n+1})$ чередуются. □

Теорема 5. *Для того, чтобы полином $P_n(x)$ степени n был полиномом наилучшего приближения функции $f(x) \in C[\,a,\ b\,]$, т.е.*

$$||f(x)\ -\ P_n(x)||_{C[a,b]}\ =\ \min_{Q_n(x) \in B_{n+1}} ||x^{n+1}\ -\ Q_n(x)||_{C[a,b]}\,,$$

где B_{n+1} – линейная оболочка функций $\{\,1,\ x,\ \ldots,\ x^n\,\}$, необходимо и достаточно, чтобы погрешность приближения $R_n(x) = f(x) - P_n(x)$ имела чебышевский альтернанс на интервале $[\,a,\ b\,]$.

Доказательство. Сначала докажем достаточность условия теоремы. Пусть

$$a\ \leq\ y_0\ <\ y_1\ <\ \ldots\ <\ y_{n+1}\ \leq\ b$$

точки чебышевского альтернанса разности $f(x) - P_n(x)$, в которых

$$|\,f(y_i) - P_n(y_i)\,|\ =\ \max_{x \in [a,b]} |\,f(x) - P_n(x)\,|\ =\ L\,,$$

знаки разности чередуются, но полином $P_n(x)$ не является полиномом наилучшего приближения, т.е. существует многочлен $Q_n(x)$ такой, что

$$|\,f(x) - Q_n(x)\,|\ <\ L\,, \qquad \forall\ x \in [\,a,\ b\,]\,.$$

Рассмотрим разность

$$Q_n(x) - P_n(x)\ =\ [\,f(x) - P_n(x)\,]\ -\ [\,f(x) - Q_n(x)\,]\,.$$

Очевидно, что знак этой разности в точках альтернанса определяется знаком первого слагаемого. Следовательно, в концах интервалов $(y_0,\ y_1),\ (y_1,\ y_2),\ \ldots,$ $(y_n,\ y_{n+1})$ функция $Q_n(x) - P_n(x)$ имеет различные знаки и, следовательно, внутри них имеет хотя бы по одному корню по теореме Ролля. Значит многочлен $Q_n(x) - P_n(x)$ степени n имеет не менее $n+1$ корня, что противоречит основной теореме алгебры. Следовательно, полином $P_n(x)$ является полиномом наилучшего приближения.

Теперь докажем необходимость условия теоремы. Пусть $P_n(x)$ является полиномом наилучшего приближения функции $f(x)$ на отрезке $[\,a,\ b\,]$. Предположим, что требуемых по условию теоремы точек альтернанса меньше: $a \leq\ y_0, < y_1 < \ldots < y_m\ \leq\ b\,,\quad 0 \leq m < n+1$, в которых $|f(y_i) - P_n(y_i)| = L$. Заметим, что хотя бы одна точка альтернанса существует, так как функция $f(x) - P_n(x)$ должна достигать своего максимуму на компакте $[\,a,\ b\,]$.

Покажем, что при таком предположении $P_n(x)$ не является полиномом наилучшего приближения функции $f(x)$.

Для определенности будем считать, что $f(y_0) - P_n(y_0) = L$.

Точек альтернанса всегда не менее двух. Действительно, если $m = 0$, то $R_n(x) > -L$ во всех точках компакта $[\,a,\ b\,]$ и, так как непрерывная функция достигает на компакте своего инфинума, то он строго больше $-L$, т.е. существует $\varepsilon > 0$ такой, что

$$-L+\varepsilon \le f(x)-P_n(x) \le +L \qquad \forall\ x\in[\,a,\ b\,].$$

или
$$-L+\frac{\varepsilon}{2} \le f(x)-[\,P_n(x)+\frac{\varepsilon}{2}\,] \le +L-\frac{\varepsilon}{2} \qquad \forall\ x\in[\,a,\ b\,],$$
т.е. полином в квадратных скобках лучше приближает функцию.

У нас имеется $m\le n$ интервалов

$$(y_0,\,y_1),\ \ (y_1,\,y_2),\ \ \ldots,\ \ (y_{m-1},\,y_m)\,.$$

Сейчас мы выберем по одной точке z_i из каждого интервала и построим полином
$$Q_{m,\varepsilon}(x) = \varepsilon\cdot(z_1-x)\cdot(z_2-x)\cdot\ldots\cdot(z_m-x)$$
такой, что разность $P_n(x)-Q_{m,\varepsilon}(x)$ будет лучше приближать функцию $f(x)$, чем $P_n(x)$.

Сначала определим вспомогательные точки

$$a=y_0^- \le y_0^+ < y_1^- \le y_1^+ < \ldots < y_{m-1}^- \le y_{m-1}^+ < y_m^- \le y_m^+ = b\,,$$

взяв в качестве y_i^+ $(i=0,\ldots,m-1)$ максимальную точку из полуинтервала $[y_i,\,y_{i+1})$, в которой значение ошибки $R_n(x)$ с совпадает с экстремальным ее значением в точке y_i;
взяв в качестве y_i^- $(i=1,\ldots,m)$ минимальную точку из интервала $(y_{i-1},\,y_i]$, в которой значение ошибки аппроксимации $R_n(x)$ с совпадает с экстремальным ее значением в точке y_i.

Тогда на компакте $[y_0^-,\,y_0^+]$

$$-L < f(x)-P_n(x) \le +L\,.$$

Внутри интервала $(y_0^+,\,y_1^-)$

$$-L < f(x)-P_n(x) < +L\,,$$

причем в концах интервала достигаются экстремальные значения (в левом: $+L$, в правом: $-L$).

Одно из решений уравнения $f(x)-P_n(x)=0$ на интервале $(y_0^+,\,y_1^-)$ возьмем в качестве z_1 (можно взять любую точку этого интервала).

Далее, на компакте $[y_1^-,\,y_1^+]$

$$-L \le f(x)-P_n(x) < +L\,.$$

Внутри интервала $(y_1^+,\,y_2^-)$

$$-L < f(x)-P_n(x) < +L\,,$$

причем в концах интервала достигаются экстремальные значения (в левом: $-L$, в правом: $+L$).

Одно из решений уравнения $f(x)-P_n(x)=0$ на интервале $(y_1^+,\,y_2^-)$ возьмем в качестве z_2 (можно взять любую точку этого интервала).

Аналогично, на компакте $[y_i^-,\,y_i^+]$, $(i=0,\ldots,m)$,
для четных i:

$$-L \; < \; f(x) - P_n(x) \; \leq \; +L \,,$$

для нечетных i :

$$-L \; \leq \; f(x) - P_n(x) \; < \; +L \,.$$

Внутри интервала $(y_i^+,\, y_{i+1}^-)$, $(i = 0, ..., m-1)$,

$$-L \; < \; f(x) - P_n(x) \; < \; +L \,,$$

причем в концах интервала достигаются экстремальные значения (в левом: $+(-1)^i L$, в правом: $-(-1)^i L)$.

Одно из решений уравнения $f(x) - P_n(x) = 0$ на интервале $(y_i^+,\, y_{i+1}^-)$ возьмем в качестве z_{i+1} (можно взять любую точку этого интервала).

При таком выборе точек $z_1 < z_2 < ... < z_m$ легко проверить, что $f(z_i) - P_n(z_i) = 0$, $(i = 1, ..., m)$, и существуют положительные числа $\varepsilon_i > 0$, $(i = 0, 1, ..., m)$, такие, что

$$\begin{array}{ll}
\forall\, x \in [a,\, z_1] & -L + \varepsilon_0 \leq f(x) - P_n(x) \; \leq \; +L \,, \\
\forall\, x \in [z_1,\, z_2] & -L \leq f(x) - P_n(x) \; \leq \; +L - \varepsilon_1 \,, \\
\quad \cdots\cdots & \qquad\qquad \cdots\cdots \\
\forall\, x \in [z_{2i},\, z_{2i+1}] & -L - \varepsilon_{2i} \leq f(x) - P_n(x) \; \leq \; +L \,, \\
\forall\, x \in [z_{2i+1},\, z_{2i+2}] & -L \leq f(x) - P_n(x) \; \leq \; +L - \varepsilon_{2i+1} \,, \\
\quad \cdots\cdots & \qquad\qquad \cdots\cdots \\
\forall\, x \in [z_{m-1},\, z_m] & -L \leq f(x) - P_n(x) \; \leq \; +L - \varepsilon_{m-1} \,, \\
\forall\, x \in [z_m,\, b] & -L + \varepsilon_m \leq f(x) - P_n(x) \; \leq \; +L
\end{array}$$

(здесь два последних неравенства для случая четного m).

Определим

$$\varepsilon \; = \; \frac{\min\limits_{0<i<m} \varepsilon_i}{2 \max\limits_{a \leq x \leq b} |(z_1 - x) \cdot (z_2 - x) \cdot ... \cdot (z_m - x)|} \,,$$

$$Q_{m,\varepsilon}(x) \; = \; \varepsilon \cdot (z_1 - x) \cdot (z_2 - x) \cdot ... \cdot (z_m - x) \,,$$

$$||Q_{m,\varepsilon}(x)||_{C[a,b]} \; = \; \min_{0<i<m} \varepsilon_i / 2 \,.$$

Практически очевидно, что многочлен $Q_{m,\varepsilon}(x)$ строго положителен на интервале $[\, a, \; z_1\,)$, строго отрицателен на интервале $(z_1,\, z_2)$ и так далее, т.е. этот полином знакоопределен на интервалах $(z_i,\, z_{i+1})$, $(i = 1, ..., m-1)$, и $(\, z_m, \; b\,]$ и меняет знак при переходе через свой корень z_i .

Результат его вычитания из разности $f(x) - P_n(x)$, как легко в этом убедиться, будет меньше $+L$ в любой точке интервала $[\, a, \; z_1\,)$, так как при каждом x вычиталось положительное число, но будет больше $-L$, так как минимальное значение разности $f(x) - P_n(x)$ на этом интервале не менее, чем $-L + \varepsilon_0$, а мы вычитаем не более, чем $\varepsilon_0 / 2$.

Аналогичными рассуждениями легко убедиться в том, что $f(x) - P_n(x) - Q_{m,\varepsilon}(x)$ строго меньше $+L$ и больше $-L$ на интервалах $(\, z_i, \; z_{i+1}\,)$, $(i =$

$1, ..., m-1)$, и $(z_m,\ b]$ и, учитывая, что эта разность обращается в нуль в точках z_i , приходим к выводу, что полином $P_n(x)+Q_{m,\varepsilon}(x)$ лучше приближает функцию $f(x)$, чем многочлен $P_n(x)$, значит m не может быть меньше $n+1$. □

3.3. Многочлен Чебышева

Найдем в явном виде решение задачи о наилучшем приближении функции $f(x) = x^{n+1}$ полиномом $P_n(x)$ степени n на отрезке $[-1,\ 1]$:

$$||x^{n+1} - P_n(x)||_{C[-1,1]} = \inf_{Q_n(x)} ||x^{n+1} - Q_n(x)||_{C[-1,1]} = L \tag{6}$$

и проверим, что корни многочлена $\omega(x) = x^{n+1} - P_n(x)$ попарно различны и принадлежат интервалу $[-1,\ 1]$.

Пусть

$$-1 \le y_0 < y_1 < ... < y_{n+1} \le 1$$

точки чебышевского альтернанса многочлена $\omega(x) = x^{n+1} - P_n(x)$, тогда эти точки являются корнями полинома $L^2 - \omega(x)^2$ степени $2(n+1)$. Так как производная этой разности обращается в нуль в точках экстремума $\omega(x)$:

$$-1 < y_1 < y_2 < ... < y_n < 1 ,$$

то последние являются корнями кратности 2, откуда следует, что корни y_0 и y_{n+1} являются простыми и производная от $\omega(x)$ в них не равна 0 . Значит точки альтернанса функции $\omega(x)$ являются корнями следующих многочленов:

$$L^2 - \omega(x)^2 = (1-x^2)\,(x-y_1)^2 ... (x-y_n)^2 ,$$
$$(1-x^2)\,[\omega(x)']^2 = (1-x^2)\,(n+1)^2\,(x-y_1)^2 ... (x-y_n)^2$$

.

Следовательно, $\omega(x)$ удовлетворяет дифференциальному уравнению

$$(1-x^2)\,[\omega(x)']^2 = (n+1)^2\,[L^2 - \omega(x)^2] ,$$

одним из решений которого на интервале $[-1,\ 1]$ является

$$\omega(x) = L\cos(\,(n+1)\,\arccos x\,) , \quad L = 2^{-n} , \tag{7}$$

корни которого, как легко проверить, задаются формулой

$$x_k = \cos\frac{(2k+1)\,\pi}{2(n+1)} , \quad k = 0, 1, ..., n , \tag{8}$$

попарно различны и принадлежат интервалу $[-1,\ 1]$.

Осталось проверить, что правая часть формулы (7) действительно является полиномом степени $(n+1)$ с коэффициентом равным единице при старшей степени аргумента. В самом деле, если $T_m(x) = \cos(\,m\,\arccos x\,)$, то

$$
\begin{aligned}
T_0(x) &= \cos(\,0\,) \;=\; 1\,, \\
T_1(x) &= \cos(\,\arccos x\,) \;=\, x\,, \\
T_2(x) &= \cos(\,2\,\arccos x\,) \;=\, 2\,x^2 - 1\,, \\
&\quad \cdots\cdots \\
T_m(x) &= 2\,x\,T_{m-1}(x) - T_{m-2}(x) \;= \\
&= 2^{m-1}\,x^m \;+\; \ldots,
\end{aligned}
\tag{9}
$$

так как

$$
\begin{aligned}
&\cos(\,m\,\arccos x\,) + \cos(\,(m-2)\,\arccos x\,) \;= \\
&\qquad\qquad = 2\,\cos(\,\arccos x\,)\,\cos(\,(m-1)\,\arccos x\,)\,.
\end{aligned}
$$

Из формул (9) следует справедливость представления (7) полинома $\omega(x)$. $\Box$

Теперь легко выписать полином $\omega(x)$ наименее уклоняющийся от нуля на отрезке $[\,a,\;b\,]$:

$$
\begin{aligned}
&\omega(x) = \Big(\frac{b-a}{2}\Big)^{n+1} \cdot \frac{1}{2^n} \cdot T_{n+1}\Big(\frac{2x-b-a}{b-a}\Big) \\
&\omega(x_k) = 0\,, \quad \frac{2x_k-b-a}{b-a} = \cos\,\frac{(2k+1)\,\pi}{2(n+1)}\,, \quad k = 0, 1, ..., n\,,
\end{aligned}
\tag{10}
$$

где $T_{n+1}(\frac{2x-b-a}{b-a})$ – полином степени $n+1$, определяемый по формулам (9).

Нетрудно проверить, что коэффициент при x^{n+1} равен единице и полином имеет $n+2$ точки альтернанса x_k :

$$
\frac{2x_k-b-a}{b-a} = \cos\,\frac{k\,\pi}{n+1}\,, \quad k = 0, 1, ..., n+1\,,
$$

что является достаточным условием наилучшего приближения x^{n+1} на отрезке $[\,a,\;b\,]$.

4. Численное интегрирование

Интегрирование функций является одной из основных математических операций. К вычислению определенных интегралов сводятся задачи определения, например, объема, площади поверхности, моментов инерции твердого тела.

В этом разделе мы рассмотрим классический подход к построению интерполяционных квадратурных формул для приближенного вычисления определенного интеграла

$$I = \int_a^b f(x)\, p(x)\, dx\,, \quad f(x) \in C\,[a,\, b]\,, \tag{1}$$

где $p(x)$ – положительная, заданная на $[a,\, b]$ функция такая, что интегралы $\int_a^b x^k\, p(x)\, dx$ существуют и конечны для любого целого $k > 0$.

Определение 1. Формула

$$I_n = \sum_{k=0}^{n} A_k\, f(x_k) \tag{2}$$

называется квадратурной формулой для приближенного вычисления интеграла (1) от функции $f(x)$ на $(n+1)$-ом узле $a \le x_0 < x_1 < \ldots < x_n \le b$ с весами $A_0,\, A_1,\, \ldots,\, A_n$. □

Так как по значением функции $f(x)$ в узлах квадратурной формулы мы можем построить интерполирующий ее полином в форме Лагранжа

$$P_n(x) = \sum_{k=0}^{n} \frac{\omega(x)}{(x - x_k)\cdot \omega'(x_k)} f(x_k)\,,$$

где

$$\omega(x) = (x - x_0)\cdot(x - x_1)\cdot \ldots \cdot (x - x_n) \tag{3}$$

– полином степени $n+1$, то естественным было бы определить веса квадратурной формулы в виде

$$A_k = \int_a^b \frac{\omega(x)}{(x - x_k)\cdot \omega'(x_k)}\, p(x)\, dx\,, \quad k = 0, 1, \ldots, n\,, \tag{4}$$

а саму формулу назвать интерполяционной квадратурной формулой.

Определение 2. Квадратурная формула (2) называется интерполяционной квадратурной формулой, если ее веса вычислются по формуле (4). □

Практически очевидно, что интерполяционная квадратурная формула (2) будет точна (т.е. $I = I_n$), если функция $f(x)$ является полиномом степени не более, чем n. И обратно, если квадратурная формула (2) точна на любом полиноме степени n, то она должна быть интерполяционной, т.е. для ее весов должна быть справедлива формула (4), что легко проверить, взяв

$$f(x) = \frac{\omega(x)}{(x - x_k) \cdot \omega'(x_k)}$$

при любом $k = 0, 1, ..., n$.

Определение 3. Алгебраической степенью точности квадратурной формулы (2) называется целое число m такое, что квадратурная формула точна на всех полиномах степени m и меньше.

Теорема 1. *Квадратурная формула (2) на $(n+1)$-ом узле будет интерполяционной квадратурной формулой тогда и только тогда, если она имеет алгебраическую степень точности n.* □

Теорема 2. *Если функция $f(x)$ $n+1$ раз непрерывно дифференцируема на отрезке $[a,\ b]$, то точность вычисления интеграла (1) по интерполяционной квадратурной формуле (2) на $(n+1)$-ом узле оценивается следующим неравенством:*

$$|I - I_n| \le \frac{M_{n+1}}{(n+1)!} \int_a^b |\omega(x)|\, p(x)\, dx\,, \tag{5}$$

где $M_{n+1} = ||f^{(n+1)}(x)||_{C[a,b]}$.

Доказательство. Теорема (1.4) утверждает, что погрешность интерполирования функции $f(x)$ по ее значениям $f(x_i)$ в попарно различных точках $x_0, x_1, ..., x_n$ интервала $[a, b]$, интерполяционным полиномом $P_n(x)$ может быть представлена в следующем виде:

$$R_n(x) = f(x) - P_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}\, \omega(x)\,,$$

где $\xi \in [a,\ b]$ зависит от x. Отсюда, учитывая, что квадратурная формула по условию теоремы является интерполяционной, следует, что

$$|I - I_n| = |\int_a^b (f(x) - P_n(x))\, p(x)\, dx| \le \frac{M_{n+1}}{(n+1)!} \int_a^b |\omega(x)|\, p(x)\, dx\,,$$

что и требовалось доказать. □

4.1. Квадратуры Гаусса наивысшей алгебраической степени точности

Итак, на любой системе узлов $a \le x_0 < x_1 < ... < x_n \le b$ интерполяционная квадратурная формула (2) точна на линейном пространстве размерности $n+1$ всех полиномов степени не более, чем n. Естественным образом возникает вопрос: а нельзя ли за счет выбора этих узлов ($n+1$ параметр) увеличить алгебраическую точность квадратурной формулы?

Теорема 3. *Квадратурная формула (2) на $(n+1)$-ом узле не может иметь алгебраическую степень точности $m > 2n+1$.*

Доказательство. Пусть $f(x) = \omega^2(x)$ – полином степени $2n+2 \le m$. Тогда

$$\sum_{k=0}^{n} A_k\, \omega^2(x_k) \;=\; 0\,,$$

так как узлы квадратуры являются корнями полинома $\omega(x)$. С другой стороны, функция $f(x) = \omega^2(x)$ неотрицательна на $[a,\, b]$, непрерывна и отлична от нуля почти всюду. Следовательно, интеграл от нее по интервалу $[a,\, b]$ строго положителен:

$$I \;=\; \int\limits_a^b \omega^2(x_k)\, p(x)\, dx \;>\; 0 \;=\; I_n\,,$$

т.е. квадратурная формула не точна. □

Теорема 4. *Если квадратурная формула (2) на $(n+1)$-ом узле $a \le x_0 < x_1 < ... < x_n \le b$ имеет алгебраическую степень точности $m = 2n+1$, то она интерполяционная, а полином $\omega(x) = (x-x_0)(x-x_1)...(x-x_n)$ ортогонален с весом $p(x)$ на $[a,\, b]$ всем полиномам меньшей степени, т.е.*

$$\int\limits_a^b \omega(x)\, P_i(x)\, p(x)\, dx \;=\; 0\,,$$

где $P_i(x)$ – произвольный полином степени $i \le n$.

Доказательство. Так как квадратурная формула (2) имеет алгебраическую степень точности $m = 2n+1 > n$, то она интерполяционная по теореме 1. Пусть $f(x) = \omega(x)\, P_i(x)$, где $P_i(x)$ – произвольный полином степени $i \le n$. Так как степень полинома $f(x)$ не превышает $2n+1$, то

$$\int\limits_a^b \omega(x)\, P_i(x)\, p(x)\, dx \;=\; \sum_{k=0}^{n} A_k\, \omega(x_k)\, P_i(x_k) \;=\; 0\,.$$

□

Теорема 5. *Если полином $\omega(x) = (x - x_0)(x - x_1)...(x - x_n)$ ортогонален с весом $p(x)$ на $[a,\, b]$ всем полиномам меньшей степени, т.е.*

$$\int_a^b \omega(x)\, P_i(x)\, p(x)\, dx \,=\, 0\,,$$

где $P_i(x)$ – произвольный полином степени $i \le n$, то интерполяционная квадратурная формула (2) на $(n+1)$-ом узле $a \le x_0 < x_1 < ... < x_n \le b$ имеет алгебраическую степень точности $m = 2n + 1$.

Доказательство. Так как квадратурная формула (2) интерполяционная, то по теореме 1 она точна на полиномах до степени n. Пусть $f(x)$ – произвольный полином степени $n+1+i \le 2n+1$. Тогда он может быть представлен в виде $f(x) = \omega(x)\, P_i(x) + Q_l(x)$, где $P_i(x)$ – некоторый полином степени $i \le n$, а полином $Q_l(x)$ степени $l \le n$ – остаток от деления $f(x)$ на $\omega(x)$. Так как полином $\omega(x)$ ортогонален с весом $p(x)$ на $[a,\, b]$ полиному $P_i(x)$, то

$$\begin{aligned} I \,=\, \int_a^b f(x)\, p(x)\, dx &= \int_a^b \omega(x)\, P_i(x)\, p(x)\, dx \,+\, \int_a^b Q_l(x)\, p(x)\, dx \\ &= \int_a^b Q_l(x)\, p(x)\, dx\,. \end{aligned}$$

С другой стороны,

$$\begin{aligned} I_n \,=\, \sum_{k=0}^{n} A_k\, f(x_k) &= \sum_{k=0}^{n} A_k\, \omega(x_k)\, P_i(x_k) \,+\, \sum_{k=0}^{n} A_k\, Q_l(x_k) \\ &= \sum_{k=0}^{n} A_k\, Q_l(x_k) \,=\, \int_a^b Q_l(x)\, p(x)\, dx \,=\, I\,, \end{aligned}$$

что и требовалось доказать. □

Теорема 6. *Полином $\omega(x) = x^{n+1} + a_n x^n + ... + a_1 x + a_0$, ортогональный с весом $p(x)$ на $[a,\, b]$ всем полиномам меньшей степени, т.е.*

$$\int_a^b \omega(x)\, P_i(x)\, p(x)\, dx \,=\, 0\,,$$

где $P_i(x)$ – произвольный полином степени $i \le n$, существует, единственнен и имеет $n+1$ простых корней на $[a,\, b]$, т.е. $\omega(x) = (x - x_0)(x - x_1)...(x - x_n)$, $a \le x_0 < x_1 < ... < x_n \le b$.

Доказательство. Выбирая $P_i(x) = x^i$, получим систему линейных алгебраических уравнений

$$\int_a^b (x^{n+1} + a_n x^n + ... + a_1 x + a_0)\, x^i\, p(x)\, dx \;=\; 0\,, \qquad i = 0,\, 1,\, ...,\, n\,. \tag{6}$$

относительно коэффициентов $a_0,\, a_1,\, ...,\, a_n$ полинома $\omega(x) = x^{n+1} + a_n x^n + ... + a_1 x + a_0$, ортогонального с весом $p(x)$ на $[a,\, b]$ всем полиномам меньшей степени.

Для существования и единственности ее решения необходимо и достаточно, чтобы однородная система

$$\int_a^b (a_n x^n + ... + a_1 x + a_0)\, x^i\, p(x)\, dx \;=\; 0\,, \qquad i = 0,\, 1,\, ...,\, n\,.$$

имела только нулевое решение. Предположим, что однородная система имеет решение $a_0,\, a_1,\, ...,\, a_n$. Тогда, умножив ее i-тое уравнение на a_i и сложив результаты, получим

$$\sum_{i=0}^{n} \int_a^b (a_n x^n + ... + a_1 x + a_0)\, a_i x^i\, p(x)\, dx \;= \;=\int_a^b (a_n x^n + ... + a_1 x + a_0)^2\, p(x)\, dx \;=\; 0\,,$$

откуда следует, что все $a_0,\, a_1,\, ...,\, a_n$ должны быть равны нулю.

Осталось исследовать корни полинома $\omega(x)$. Если этот полином имеет $n + 1$ корень нечетной кратности (т.е. простые корни) на $[a,\, b]$ то теорема доказана.

Если корней нечетной кратности нет, то определим полином $Q(x) = 1$ нулевой степени, если $a \le x_0 < x1 < ... < x_m \le b$, $m < n$, – все его корни нечетной кратности на $[a,\, b]$ то определим полином $Q(x) = (x - x_0)(x - x_1)...(x - x_m)$. Тогда произведение $\omega(x)\, Q(x)$ не меняет знак на $[a,\, b]$ и, следовательно, в силу непрерывности и отличия от нуля почти всюду интеграл от него по отрезку $[a,\, b]$ с положительным весом $p(x)$ отличен от нуля, что противоречит ортогональности полинома $\omega(x) = x^{n+1} + a_n x^n + ... + a_1 x + a_0$ с весом $p(x)$ на $[a,\, b]$ всем полиномам меньшей степени. Значит, $m = n$. $\square$

Таким образом теоремы 3-6 гарантируют существование единственной интерполяционной квадратурной формулы наивысшей алгебраической точности (квадратуры Гаусса).

Теорема 7. *Если функция $f(x)$ $2(n+1)$ раз непрерывно дифференцируема на отрезке $[\,a,\,b\,]$, то точность вычисления интеграла (1) по квадратурной формуле Гаусса (2) на $(n+1)$-ом узле оценивается следующим неравенством:*

$$|I - I_n| \leq \frac{M_{2(n+1)}}{(2(n+1))!} \int\limits_a^b |\omega^2(x)|\, p(x)\, dx\,, \tag{7}$$

где $M_{2(n+1)} = ||f^{(2n+2)}(x)||_{C[a,b]}$.

Доказательство. Из теоремы (2.3) следует, что погрешность интерполирования функции $f(x)$ по ее значениям $f(x_i)$ и значениям ее первой производной $f'(x_i)$ в попарно различных точках $x_0, x_1, ..., x_n$ интервала $[\,a,\,b\,]$ интерполяционным полиномом $P_{2n+1}(x)$ может быть представлена в следующем виде:

$$R_{2n+1}(x) = f(x) - P_{2n+1}(x) = \frac{f^{(2n+2)}(\xi)}{(2(n+1))!}\,\omega^2(x)\,,$$

где $\xi \in [\,a,\,b\,]$ зависит от x. Отсюда, учитывая, что квадратурная формула Гаусса точна на полиномах до степени $2n+1$, следует, что

$$|I - I_n| = |\int\limits_a^b (f(x) - P_{2n+1}(x))\, p(x)\, dx| \leq \frac{M_{2(n+1)}}{(2(n+1))!} \int\limits_a^b |\omega^2(x)|\, p(x)\, dx\,,$$

что и требовалось доказать. □

4.2. Сходимость квадратур Гаусса

Замечательным свойством квадратур Гаусса является их сходимость для любой функции $f(x) \in C\,[a,\,b]$.

Теорема 8. *Если функция $f(x) \in C\,[\,a,\,b\,]$, то интерполяционная квадратурная формула (2) на $(n+1)$-ом узле с положительными весами сходится к интегралу (1) при $n \to \infty$.*

Доказательство. Поскольку множество всех полиномов плотно в $C\,[\,a,\,b\,]$, то для любого сколь угодно малого $\varepsilon > 0$ существует полином $Q_N(x)$ степени N такой, что

$$|\,f(x) - Q_N(x)\,| \leq \frac{\varepsilon}{2} \quad \forall x \in [\,a,\,b\,]\,.$$

Тогда для любого $n \geq N$, любой интерполяционной квадратурной формулы (2) имеем

$$\int\limits_a^b Q_N(x)\, p(x)\, dx = \sum_{k=0}^{n} A_k\, Q_N(x_k)\,,$$

$$I - I_n = \int\limits_a^b f(x)\,p(x)\,dx - \sum_{k=0}^{n} A_k\,f(x_k) =$$

$$= \int\limits_a^b f(x)\,p(x)\,dx - \int\limits_a^b Q_N(x)\,p(x)\,dx + \sum_{k=0}^{n} A_k\,Q_N(x_k) - \sum_{k=0}^{n} A_k\,f(x_k) =$$

$$= \int\limits_a^b [f(x) - Q_N(x)]\,p(x)\,dx + \sum_{k=0}^{n} A_k\,[Q_N(x_k) - f(x_k)]\,.$$

Так как интерполяционная квадратурная формула точна на константе, а по условиям теоремы ее веса положительны, то

$$|I - I_n| \le \int\limits_a^b |f(x) - Q_N(x)|\,p(x)\,dx + \sum_{k=0}^{n} A_k\,|Q_N(x_k) - f(x_k)| \le$$

$$\le \frac{\varepsilon}{2}\int\limits_a^b p(x)\,dx + \frac{\varepsilon}{2}\sum_{k=0}^{n} A_k \; = \varepsilon \int\limits_a^b p(x)\,dx\,,$$

что и требовалось доказать. □

Осталось показать, что веса квадратур Гаусса положительны.

Лемма 1. *Все веса квадратурной формулы Гаусса (2) положительны.*

Доказательство. Пусть

$$I_n \;=\; \sum_{k=0}^{n} A_k\,f(x_k)$$

квадратура Гаусса и, следовательно, она точна на полиномах до степени $2n+1$, а значит и на полиноме

$$f(x) = \Big[\frac{\omega(x)}{(x - x_i)}\Big]^2,$$

где x_i один из узлов квадратуры (корней полинома $\omega(x)$).

Тогда, поскольку $f(x)$ неотрицательна, непрерывна и не равна тождественно нулю на $[\,a,\,b\,]$,

$$0 < I = \int\limits_a^b f(x)\,p(x)\,dx = I_n = \sum_{k=0}^{n} A_k\,f(x_k) =$$

$$= \sum_{k=0}^{n} A_k\,\Big[\frac{\omega(x_k)}{(x_k - x_i)}\Big]^2 = A_i\big[\omega'(x_i)\big]^2,$$

откуда следует положительность i-того веса квадратуры Гаусса. □

Заметим, что этот результат не имеет места для квадратурных интерполяционных формул на равномерной сетке узлов.

4.3. Устойчивость квадратурных формул

Зачастую, по тем или иным причинам, значения интегрируемой функции вычисляются неточно: вместо $f(x_k)$ мы имеем $f_\varepsilon(x_k)$. Тогда

$$|I_n - I_{n,\varepsilon}| = |\sum_{k=0}^{n} A_k\, f(x_k) - \sum_{k=0}^{n} A_k\, f_\varepsilon(x_k)| \le$$

$$\le \varepsilon \sum_{k=0}^{n} |A_k| = \varepsilon \int\limits_a^b p(x)\, dx\,,$$

если квадратура точна на константе и ее веса положительны, а $|f(x) - f_\varepsilon(x)| \le \varepsilon$ во всех квадратурных узлах. Отсюда следует, что малые изменения интегрируемой функции мало изменяют приближенное значение интеграла независимо от числа квадратурных узлов.

Заметим, что таким свойством не обладают квадратурные интерполяционные формулы на равномерной сетке узлов.

5. Простейшие квадратурные формулы

В этом разделе для приближенного вычисления определенного интеграла

$$I = \int\limits_a^b f(x)\,dx\,, \tag{1}$$

выводятся простейшие квадратурные интерполяционные формулы

$$I_n = \sum_{k=0}^{n} A_k\, f(x_k)\,, \quad A_k = \int\limits_a^b \frac{\omega(x)}{(x-x_k)\cdot\omega'(x_k)}\,dx\,, \\ \omega(x) = (x-x_0)\cdot(x-x_1)\cdot\ldots\cdot(x-x_n)\,, \tag{2}$$

на одном, двух и трех узлах; конкретизируется оценка их точности (см. (4.5))

$$|I - I_n| \le \frac{M_{n+1}}{(n+1)!}\int\limits_a^b |\omega(x)|\,dx\,, \tag{3}$$

где $M_{n+1} = ||f^{(n+1)}(x)||_{C[a,b]}$. Кроме того, в первых двух случаях строятся квадратуры Гаусса с оценкой точности (см. (4.7))

$$|I - I_n| \le \frac{M_{2(n+1)}}{(2(n+1))!}\int\limits_a^b |\omega^2(x)|\,dx\,, \tag{4}$$

Всюду через $h = b - a$ мы будем обзначать длину интервала интегрирования, а через c его среднюю точку $(a+b)/2$.

5.1. Формулы прямоугольников (на одном узле)

Сначала выберем в качестве квадратурного узла $x_0 = a$. Тогда

$$x_0 = a\,, \quad \omega(x) = x - a\,, \quad A_0 = \int\limits_a^b dx = h\,, \quad \int\limits_a^b |x-a|\,dx = \frac{(b-a)^2}{2}$$

и, следовательно,

$$I_0 = f(a)\,(b-a)\,, \quad |I - I_0| \le \frac{M_1}{2}(b-a)^2 \tag{5}$$

– формула второго порядка малости по h. □

Теперь выберем в качестве квадратурного узла $x_0 = c$. Тогда

$$x_0 = c\,,\quad \omega(x) = x - c\,,\quad A_0 = \int\limits_a^b dx = h\,,\quad \int\limits_a^b |x-c|\,dx = \frac{(b-a)^2}{4}\,.$$

Так как $\int\limits_a^b (x-c)\,dx = 0$, т.е. $\omega(x)$ ортогонален полиномам меньшей степени и мы имеем квадратуру Гаусса, то

$$I_0 \;=\; f(c)\,(b-a)\,,\quad |I - I_0| \le \frac{M_2}{24}(b-a)^3\,, \tag{6}$$

поскольку $\int\limits_a^b (x-c)^2\,dx = \dfrac{(b-a)^3}{12}$.

Следовательно, формула прямоугольников с центральной точкой – формула третьего порядка малости по h.

5.2. Формула трапеций (на двух узлах)

Выберем в качестве квадратурных узлов $x_0 = a$ и $x_1 = b$. Тогда

$$x_0 = a\,,\quad x_1 = b\,,\quad \omega(x) = (x-a)\,(x-b)\,,$$

$$A_0 = \int\limits_a^b \frac{x-b}{a-b}\,dx = \frac{h}{2}\,,\quad A_1 = \int\limits_a^b \frac{x-a}{b-a}\,dx = \frac{h}{2}\,,$$

$$\int\limits_a^b |(x-a)(x-b)|\,dx = -\int\limits_{-0.5\,h}^{0.5\,h} (y+0.5\,h)(y-0.5\,h)\,dy = \frac{(b-a)^3}{6}$$

и, следовательно,

$$I_1 \;=\; \frac{f(a)+f(b)}{2}\,(b-a)\,,\quad |I - I_1| \le \frac{M_2}{12}(b-a)^3 \tag{7}$$

– формула третьего порядка малости по h. □

Для того, чтобы получить квадратуру Гаусса на двух узлах, нужно построить полином $\omega(x) = (x-x_0)\,(x-x_1)$ ортогональный 1 и x. Будем искать его в виде $\omega(x) = (x-a)\,(x-b) + d$. Тогда параметр d определяется из уравнения

$$\int\limits_a^b [(x-a)(x-b)+d]\,dx = \int\limits_{-0.5\,h}^{0.5\,h} [(y+0.5\,h)(y-0.5\,h)+d]\,dy =$$

$$= h\Big[d - \frac{1}{6}h^2\Big] \;=\; 0\,,$$

а корни полинома $\omega(x)$ легко вычисляются: $x_0 = c - \dfrac{\sqrt{3}}{6}h$ и $x_1 = c + \dfrac{\sqrt{3}}{6}h$. Осталось проверить ортогональность полинома $\omega(x)$ функции x:

$$\int\limits_a^b [(x-a)(x-b)+d]\,x\,dx = \int\limits_{-0.5\,h}^{0.5\,h} [(y+0.5\,h)(y-0.5\,h)+d]\,(y+c)\,dy =$$

$$= \int\limits_{-0.5\,h}^{0.5\,h} \Big[(y+0.5\,h)(y-0.5\,h)+\frac{1}{6}h^2\Big]\,y\,dy \;=\; 0\,.$$

Тогда

$$x_0 = c - \frac{\sqrt{3}}{6}\,, \quad x_1 = c + \frac{\sqrt{3}}{6}\,, \quad \omega(x) = (x-c)^2 - \frac{(b-a)^2}{12}\,,$$

$$A_0 = \frac{-\sqrt{3}}{b-a}\int\limits_a^b (x-c-\frac{\sqrt{3}}{6}h)\,dx = \frac{-\sqrt{3}}{b-a}\int\limits_{-0.5\,h}^{0.5\,h} (y-\frac{\sqrt{3}}{6}h)\,dy = \frac{h}{2} = A_1\,,$$

$$\int\limits_a^b \omega^2(x)\,dx = -\int\limits_{-0.5\,h}^{0.5\,h} \Big[y^2 - \frac{h^2}{12}\Big]^2\,dy = \frac{(b-a)^5}{180}$$

и, следовательно,

$$I_1 \;=\; \Big[f(c-\frac{\sqrt{3}}{6})+f(c+\frac{\sqrt{3}}{6})\Big]\,\frac{b-a}{2}\,, \quad |I-I_1| \le \frac{M_4}{4320}(b-a)^5 \qquad (8)$$

– формула пятого порядка малости по h.

5.3. Формула Симпсона (на трех узлах)

Выберем в качестве квадратурных узлов $x_0 = a$, $x_1 = c$ и $x_2 = b$. Тогда

$$x_0 = a\,, \quad x_1 = c\,, \quad x_2 = b\,, \quad \omega(x) = (x-a)\,(x-c)\,(x-b)\,,$$

$$A_0 = \int\limits_a^b \frac{(x-c)(x-b)}{(a-c)(a-b)}\,dx = \frac{2}{h^2}\int\limits_{-0.5\,h}^{0.5\,h} y\,(y-0.5\,h)\,dy = \frac{h}{6}\,,$$

$$A_1 = \int\limits_a^b \frac{(x-a)(x-b)}{(c-a)(c-b)}\,dx = -\frac{4}{h^2}\int\limits_{-0.5\,h}^{0.5\,h} (y+0.5\,h)\,(y-0.5\,h)\,dy = \frac{4\,h}{6}\,,$$

$$A_2 = \int\limits_a^b \frac{(x-a)(x-c)}{(b-a)(b-c)}\,dx = \frac{2}{h^2}\int\limits_{-0.5\,h}^{0.5\,h} (y+0.5\,h)\,y\,dy = \frac{h}{6}\,.$$

Кроме того, полином $\omega(x)$ ортогонален единице:

$$\int\limits_a^b (x-a)(x-c)(x-b)\,dx = \int\limits_{-0.5\,h}^{0.5\,h} (y+0.5\,h)\,y\,(y-0.5\,h)\,dy = 0\,,$$

а это значит, что соответствующая квадратурная формула точна на многочленах до третьей степени. Действительно, если $f(x)$ – полином третьей степени, то имеет место представление

$$f(x) = A\,\omega(x) + Q_2(x)$$

с некоторыми постоянной A и полиномом второй степени $Q_2(x)$. Тогда

$$I = \int\limits_a^b f(x)\,dx = \int\limits_a^b Q_2(x)\,dx = \sum_{k=0}^{2} A_k\,Q_2(x_k) = \sum_{k=0}^{2} A_k\,f(x_k) = I_2\,.$$

Выбирая в качестве интерполянта для функции $f(x)$ полином Эрмита P_3 на двух простых узлах x_0, x_2 и узле x_1 кратности два, получим

$$|f(x) - P_3(x)| \le \frac{M_4}{24}\,|(x-a)\,(x-c)^2\,(x-b)|\,,$$

$$\int\limits_a^b P_3(x)\,dx = \sum_{k=0}^{2} A_k\,f(x_k) = I_2$$

$$|I - I_2| \le \frac{M_4}{24}\int\limits_a^b |(x-a)\,(x-c)^2\,(x-b)|\,dx\,.$$

Так как

$$\int\limits_a^b |(x-a)\,(x-c)^2\,(x-b)|\,dx = -\int\limits_{-0.5\,h}^{0.5\,h} (y+0.5\,h)\,y^2\,(y-0.5\,h)\,dy = \frac{h^5}{120}\,,$$

то квадратурная формула Симпсона имеет пятый порядок по h:

$$I_2 = \frac{f(a) + 4f(c) + f(b)}{6}\,(b-a)\,, \quad |I - I_2| \le \frac{M_4}{2880}(b-a)^5\,. \tag{9}$$

5.4. Составные квадратурные формулы

Использование простейших квадратурных формул для интегрирования функции на интервале, длина которого не является малой величиной, редко приводит к хорошим результатам. Поэтому отрезок интегрирования обычно разбивают на на подъинтервалы малой длины, исходный интеграл представляют в виде суммы интегралов по отрезкам разбиения и каждый интеграл этой суммы заменяют по той или иной простейшей квадратурной формуле.

Вернемся к задаче приближенного вычисления определенного интеграла

$$I \;=\; \int\limits_a^b f(x)\,dx\,.$$

Будем считать, что интервал $[\,a,\,b\,]$ разбит на N подъинтервалов $[\,a_k,\,b_k\,]$ одинаковой длины $h = (b-a)/N$:

$$[\,a_k,\,b_k\,] \;=\; [\,a + (k-1)h,\,a + kh\,]\,, \quad k = 1, 2, ..., N\,.$$

Через $M_{n,k}$ будем обозначать максимум модуля n-той производной функции $f(x)$ на интервале $[\,a_k,\,b_k\,]$, а через c_k среднюю точку этого интервала.

5.4.1. Составные формулы прямоугольников

Применяя формулу прямоугольников (5), получим, что

$$I \approx I_h = \sum_{k=1}^{N} f(a_k)\, h\,, \quad |I - I_h| \le \sum_{k=1}^{N} \frac{M_{1,k}}{2} h^2 \le \frac{M_1}{2} h \tag{10}$$

– формула первого порядка малости по h. □

Применяя формулу прямоугольников (6), получим, что

$$I \approx I_h = \sum_{k=1}^{N} f(c_k)\, h\,, \quad |I - I_h| \le \sum_{k=1}^{N} \frac{M_{2,k}}{24} h^3 \le \frac{M_2}{24} h^2 \tag{11}$$

– формула второго порядка малости по h. □

5.4.2. Составная формула трапеций

Применяя формулу трапеций (7), получим, что

$$\begin{gathered} I \approx I_h = \sum_{k=1}^{N} \frac{f(a_k) + f(b_k)}{2}\, h\,, \\ |I - I_h| \le \sum_{k=1}^{N} \frac{M_{2,k}}{12} h^3 \le \frac{M_2}{12} h^2 \end{gathered} \tag{12}$$

– формула второго порядка малости по h.

5.4.3. Составная формула Симпсона

Применяя формулу Симпсона (9), получим, что

$$\begin{gathered} I \approx I_h = \sum_{k=1}^{N} \frac{f(a_k) + 4f(c_k) + f(b_k)}{6}\, h\,, \\ |I - I_h| \le \sum_{k=1}^{N} \frac{M_{4,k}}{2880} h^5 \le \frac{M_4}{2880} h^4 \end{gathered} \tag{13}$$

– формула четвертого порядка малости по h.

6. Итерационные методы решения нелинейных числовых уравнений

Пусть на отрезке $[\,a,\,b\,]$ задана непрерывная функция $f(x)$ и требуется решить уравнение

$$f(x) \,=\, 0\,. \tag{1}$$

Любое решение $x^* \in [\,a,\,b\,]$ этого уравнения будем называть корнем (нулем) функции $f(x)$. Отметим, что каких-либо общих правил анализа расположение корней произвольной непрерывной функции $f(x)$ на отрезке $[\,a,\,b\,]$ не существует.

Определение 1. Корень x^* функции $f(x)$ будем называть изолированным корнем кратности $p > 0$, если

$$f(x) \,=\, (x - x^*)^p\, g(x) \tag{2}$$

и функция $g(x)$ непрерывна на отрезке $[\,a,\,b\,]$ и знакоопределена в некоторой окрестности $I_\varepsilon(x^*) = [\,x^* - \varepsilon,\, x^* + \varepsilon\,] \cap [\,a,\,b\,]$, $\varepsilon > 0$. □

Очевидно, что, найдя один изолированный корень функции $f(x)$ и определив его кратность, следующий корень мы можем искать как корень функции $g(x)$.

Построить явные формулы, определяющие корни непрерывной функции, удается крайне редко, поэтому основными методами решения задачи (1) являются итерационные процесссы, позволяющие получить ее решение с любой точностью. Мы ограничимся изучением двух методов: метода последовательных приближений и метода Ньютона, основой которых является принцип сжимающих отображений.

6.1. Принцип сжимающих отображений

Пусть M – метрическое пространство (любое нормированное пространство является метрическим) с расстоянием (метрикой) $\rho(x, y)$ между его элементами x и y.

Определение 2. Отображение

$$\varphi\,:\, M \to M \tag{3}$$

называется сжимающим, если существует положительное число $q < 1$ такое, что

$$\rho(\varphi(x), \varphi(y)) \le q\,\rho(x, y)\,, \quad \forall\, x, y \in \,. \tag{4}$$

Определение 3. Элемент $x \in M$ называется неподвижной точкой отображения φ из M в M, если

$$x \,=\, \varphi(x)\,. \tag{5}$$

Теорема 1 (принцип сжимающих отбражений).
Любое сжимающее отображение $\varphi(x)$ из M в M, где M – полное метрическое пространство, имеет одну и только одну неподвижную точку $x^* \in M$. Более того, последовательность (метод последовательных приближений или простой итерации)

$$x_{n+1} = \varphi(x_n)\,, \quad x_0 \in M\,, \quad n = 0,\, 1,\, \ldots, \tag{6}$$

сходится к этой точке и имеет место оценка

$$\rho(x^*, x_n) \le \frac{\rho(x_1, x_0)}{1-q}\, q^n\,. \tag{7}$$

Доказательство. Пусть x_0 – произвольный элемент (точка) из M. Докажем, что последовательность (6) имеет предел. Для этого достаточно, показать, что она фундаментальна по Коши.

Из определений последовательности и сжимающего отображения следует, что

$$\begin{aligned}\rho(x_{n+m}, x_n) &= q\,\rho(\varphi(x_{n+m-1}), \varphi(x_{n-1}) \le \\ &\le q\,\rho(x_{n+m-1}, x_{n-1}) \le\, \ldots\, \le q^n\,\rho(x_m, x_0)\,.\end{aligned}$$

Из этого неравенства (и неравенства треугольника) следует, что

$$\begin{aligned}\rho(x_m, x_0) &\le \rho(x_m, x_{m-1}) + \rho(x_{m-1}, x_0) \le \\ &\le q^{m-1}\,\rho(x_1, x_0) + \rho(x_{m-1}, x_0) \le, \ldots \le \\ &\le (q^{m-1} + q^{m-2} + \ldots + q + 1)\,\rho(x_1, x_0) \le \\ &\le \frac{\rho(x_1, x_0)}{1-q}\end{aligned}$$

и, значит,

$$\rho(x_{n+m}, x_n) \le \frac{\rho(x_1, x_0)}{1-q}\, q^n\,. \tag{8}$$

Так как $q < 1$, то из (8) следует фундаментальность по Коши последовательности (6), а из полноты метрического пространства M ее сходимость к некоторому элементу $x^* = \varphi(x^*) \in M$, т.е. этот предел является неподвижной точкой сжимающего (а значит и непрерывного) отображения $\varphi(x)$.

Более того, переходя в неравенстве (8) к пределу при $m \to \infty$, получим оценку сходимости (7).

Единственность неподвижной точки устанавливается элементарно: пусть x^* и x^{**} – два решения уравнения (5). Тогда

$$\rho(x^*, x^{**}) = \rho(\varphi(x^*), \varphi(x^{**})) \le q\,\rho(x^*, x^{**})$$

и, так как $q < 1$, то $x^* = x^{**}$. □

6.2. Метод простой итерации

Вернемся к уравнению (1) на отрезке $[\,a,\,b\,]$ и преобразуем его к виду

$$x \;=\; \varphi(x)\,, \tag{9}$$

задав, например, $\varphi(x) = x - \tau(x)\,f(x)$, где $\tau(x)$ – непрерывная, знакоопределенная на отрезке $[\,a,\,b\,]$ функция.

Рассмотрим для решения уравнения (9) метод простой итерации (6):

$$x_{n+1} \;=\; \varphi(x_n)\,, \quad x_0 \in [\,a,\,b\,]\,, \quad n \;=\; 0,\,1,\,\ldots\,, \tag{10}$$

Теорема 2. *Если функция $\varphi(x)$ на отрезке $[\,a,\,b\,]$ непрерывна по Липшицу:*

$$|\,\varphi(x) - \varphi(y)\,| \le q\,|\,x - y\,|\,, \quad q \;<\; 1\,, \tag{11}$$

и преобразует интервал $[\,a,\,b\,]$ в $[\,a,\,b\,]$, т.е.

$$a \;\le\; \varphi(x) \;\le\; b \qquad \forall\, x \in [\,a,\,b\,]\,, \tag{12}$$

то уравнение (9) имеет единственное на $[\,a,\,b\,]$ решение x^*, к которому сходится последовательность $\{x_n\}$ метода простой итерации (10) при любом начальном приближении $x_0 \in [\,a,\,b\,]$ и имеет место оценка

$$|\,x^* - x_n\,| \;\le\; \frac{|\,x_1 - x_0\,|}{1 - q}\,q^n\,. \tag{13}$$

Доказательство. Определим метрическое пространство $M = [\,a,\,b\,]$ с расстоянием $\rho(x, y) = |x - y|$. Очевидно, что оно полно, а из условий (11) и (12) следует, что $\varphi(x)$ – сжимающее отображение из M в M и, значит, утверждения этой теоремы являются прямым следствием теоремы (1). $\square$

Проверка условий теоремы 2 для конкретных функций $\varphi(x)$ может вызвать определенные затруднения. Действительно, условие (11) для непрерывно дифференцируемой функции означает ограниченность на интервале $[\,a,\,b\,]$ модуля ее первой производной постоянной Липшица $q < 1$, а проверка условия (12) заключается в поиске (или оценке) экстремальных значений самой функции, что может оказаться более сложной задачей по сравнению с решением уравнения (9).

Если константа Липшица $q < 1$ известна, то при достаточно удачном выборе начального приближения x_0 метода простой итерации удается избавиться от проверки условия (12).

Теорема 3. *Если функция $\varphi(x)$ на отрезке $[\,a,\,b\,]$ непрерывна по Липшицу:*

$$|\,\varphi(x) - \varphi(y)\,| \le q\,|\,x - y\,|\,, \quad q \;<\; 1\,,$$

и известно $x_0 \in [\,a,\,b\,]$ такое, что

$$[\,x_0-\delta,\,x_0+\delta\,]\subset[\,a,\,b\,]\,,\quad \delta=\frac{|\,\varphi(x_0)-x_0\,|}{1-q}\,, \tag{14}$$

то уравнение $x=\varphi(x)$ *имеет в* $[\,x_0-\delta,\,x_0+\delta\,]$ *единственное решение* x^*, *к которому сходится последовательность* $\{x_n\}$ *метода простой итерации*

$$x_{n+1}\,=\,\varphi(x_n)\,,\qquad n\,=\,0,\,1,\,\ldots,$$

и имеет место оценка

$$|\,x^*-x_n\,|\,\le\,\delta\,q^n\,.$$

Доказательство. Функция $\varphi(x)$ отображает отрезок $[\,x_0-\delta,\,x_0+\delta\,]$ в себя. Действительно, $\forall\,x\in[\,x_0-\delta,\,x_0+\delta\,]$

$$\begin{aligned}|\,\varphi(x)-x_0\,|&\le|\,\varphi(x)-\varphi(x_0)\,|+|\,\varphi(x_0)-x_0\,|\le\\&\le q\,|\,x-x_0\,|+(1-q)\,\delta\le\delta\,.\end{aligned}$$

Следовательно, выполняются все условия теоремы 2 на интервале $[\,x_0-\delta,\,x_0+\delta\,]$, откуда следует сходимость и оценка сходимости метода простой итерации.
□

Следующее утверждение характеризует кратность корня функции $f(x)=x-\varphi(x)$.

Теорема 4. *Если функция* $\varphi(x)$ *на отрезке* $[\,a,\,b\,]$ *непрерывна по Липшицу:*

$$|\,\varphi(x)-\varphi(y)\,|\le q\,|\,x-y\,|\,,\quad q\,<\,1\,,$$

и преобразует интервал $[\,a,\,b\,]$ *в* $[\,a,\,b\,]$, *т.е.*

$$a\,\le\,\varphi(x)\,\le\,b\qquad\forall\,x\in[\,a,\,b\,]\,,$$

то функция $f(x)=x-\varphi(x)$ *имеет единственный на* $[\,a,\,b\,]$ *корень* x^* *кратности равной единице.*

Доказательство. Существование и единственность корня x^* функции $f(x)$ на отрезке $[\,a,\,b\,]$ следует из теоремы 2.

Используя неравенство треугольника и непрерывность по Липшицу функции $\varphi(x)$, получим, что $\forall\,x\in[\,a,\,b\,]$

$$\begin{aligned}|\,x-x^*\,|\,-\,|\,f(x)\,|\le|\,x-x^*\,-\,f(x)\,|\,&=\,|\,\varphi(x)-\varphi(x^*)\,|\,\le\\&\le q\,|\,x-x^*\,|\end{aligned}$$

и

$$\begin{aligned}|\,f(x)\,|\,-\,|\,x-x^*\,|\le|\,f(x)\,-\,(x-x^*)\,|\,&=\,|\,-\varphi(x)\,+\,\varphi(x^*)\,|\,\le\\&\le q\,|\,x-x^*\,|\,,\end{aligned}$$

т.е.

$$0\,<\,(\,1-q\,)\,\le\,\frac{|\,f(x)\,|}{|\,x-x^*\,|}\,\le\,(\,1+q\,)\,.$$

Пример. Рассмотрим задачу решения уравнения $f(x) = x^2 - a = 0$ при $a > 1$ на интервале $[1, a]$. Перепишем ее в эквивалентной форме

$$x = \varphi(x) \equiv x - \frac{x^2 - a}{1 + a}.$$

Легко проверить, что $\forall x \in [1, a]$

$$|\varphi(x) - \varphi(y)| \le \frac{a - 1}{a + 1} = q < 1, \quad 1 < \varphi(x) < a,$$

т.е. выполняются условия теоремы 2 и, следовательно, последовательность $\{x_n\}$ метода простой итерации

$$x_{n+1} = x_n - \frac{x_n^2 - a}{1 + a}$$

сходится к квадратному корню из числа a при любом начальном приближении $x_0 \in [1, a]$.

Для погрешности $\varepsilon_n = x_n - \sqrt{a}$ легко получить следующее представление

$$\varepsilon_{n+1} = \varepsilon_n - \frac{x_n^2 - a}{1 + a} = \left(1 - \frac{x_n + \sqrt{a}}{1 + a}\right)\varepsilon_n \approx \left(1 - \frac{2\sqrt{a}}{1 + a}\right)\varepsilon_n.$$

Заметим, что сходимость (ее скорость) называют линейной, если погрешность очередного приближения пропорциональна (конечно с коэффициентом меньшим единицы) погрешности на предыдущей итерации. Следовательно, в нашем примере скорость сходимости линейна.

Метод последовательных приближений также называют методом Якоби.

6.3. Метод Эйткена ускорения сходимости

Предположим, что $x^* \in [a, b]$ – решение уравнения

$$x^* = \varphi(x^*), \tag{15}$$

к которому сходятся последовательные приближения

$$x_{n+1} = \varphi(x_n), \quad x_0 \in [a, b], \quad n = 0, 1, \ldots. \tag{16}$$

Пусть $\varphi(x)$ – дважды непрерывно дифференцируемая на $[a, b]$ функция и $\varphi'(x) = a \neq 0$. Тогда, так как

$$\varphi(x_n) = \varphi(x^*) + \varphi'(x^*)(x_n - x^*) + \frac{\varphi''(\xi_n)}{2}(x_n - x^*)^2,$$

из (15) и (16) следует, что

$$\begin{aligned} x_{n+1} - x^* &= a(x_n - x^*) + \varepsilon_n, \\ x_{n+2} - x^* &= a(x_{n+1} - x^*) + \varepsilon_{n+1}, \end{aligned}$$

где $\varepsilon_n = O(|x_n - x^*|^2)$ и $\varepsilon_{n+1} = O(|x_n - x^*|^2)$.

Разрешая эту систему относительно неизвестных x^* и a, получим

$$
\begin{aligned}
x^* = x_{n+2} - \frac{(\Delta x_{n+1})^2}{\Delta^2 x_n} + \\
+ \varepsilon_{n+1} \frac{\Delta x_n}{\Delta \varepsilon_n - \Delta^2 x_n} + \Delta \varepsilon_n \, \Delta x_{n+1} \, \frac{\Delta x_n}{\Delta^2 x_n - \Delta \varepsilon_n} \,,
\end{aligned}
\tag{17}
$$

где

$$
\Delta x_n = x_{n+1} - x_n \,, \quad \Delta x_{n+1} = x_{n+2} - x_{n+1} \,, \quad \Delta^2 x_n = \Delta x_{n+1} - \Delta x_n \,,
$$
$$
\Delta \varepsilon_n = \varepsilon_{n+1} - \varepsilon_n \,.
$$

Нетрудно показать, что

$$
\frac{\Delta x_n}{\Delta^2 x_n - \Delta \varepsilon_n} = (a-1)^{-1} + O(|x_n - x^*|) \,,
$$

откуда, учитывая равенство $\varepsilon_{n+1} = O(|x_n - x^*|^2)$, следует, что

$$
\tilde{x}_{n+2} = x_{n+2} - \frac{(\Delta x_{n+1})^2}{\Delta^2 x_n} \tag{18}
$$

отличается от x^* на величину $O(|x_n - x^*|^2)$, т.е. пропорциональна квадрату погрешности на n-том шаге итерационного процесса (квадратичная скорость сходимости).

Формула (18) построения по трем последовательным приближениям x_n, x_{n+1} и x_{n+2} значительно (на порядок) более близкого к x^* приближения $\tilde{x}_{n+2}$ называется методом Эйткена ускорения простой итерации.

Метод Эйткена позволяет предположить о возможности построения для решения уравнения $f(x) = 0$ итерационных процессов с квадратичной скоростью сходимости.

7. Метод Ньютона

Пусть на отрезке $[\,a,\,b\,]$ задана непрерывно дифференцируемая функция $f(x)$ и требуется решить уравнение

$$f(x) \,=\, 0\,. \tag{1}$$

Идея метода Ньютона чрезвычайно проста: если x_n – приближение к корню x^*, то функция $f(x)$ заменяется интерполяционным полиномом первой степени по ее значению $f(x_n)$ и значению ее первой производной $f\,'(x_n)$:

$$f(x) \,\approx\, f(x_n) \,+\, f\,'(x_n)\,,$$

а корень этого полинома объявляется следующим приближением:

$$x_{n+1} \,=\, x_n \,-\, \frac{f(x_n)}{f\,'(x_n)}\,, \quad x_0 \in [\,a,\,b\,]\,, \quad n \,=\, 0,\,1,\,\dots\,. \tag{2}$$

Очевидно, что метод Ньютона (2) решения уравнения (1) можно трактовать как метод последовательных приближений решения эквивалентного уравнения

$$x \,=\, \varphi(x) \,\equiv\, x \,-\, \frac{f(x)}{f\,'(x)}\,. \tag{3}$$

Пусть $x^* \in [\,a,\,b\,]$ – изолированный корень кратности $p > 0$ непрерывной функции $f(x)$:

$$f(x) \,=\, (x - x^*)^p\, g(x)\,, \tag{4}$$

где функция $g(x)$ дважды непрерывно дифференцируема и знакоопределена в некоторой окрестности корня.

Тогда, так как

$$\begin{aligned} \varphi(x) &= \frac{p-1}{p}x + \frac{x^*}{p} + \frac{g'(x)\,(x-x^*)^2}{p\,(\,p\,g(x) + g'(x)\,(x-x^*)\,)}\,, \\ \varphi'(x) \,&=\, \frac{p-1}{p} + O(|x-x^*|)\,, \end{aligned} \tag{5}$$

приближения x_n метода Ньютона (2) к корню x^* функции $f(x)$ удовлетворяют соотношениям

$$\begin{aligned} x_{n+1} - x^* = \varphi(x_n) - \varphi(x^*) = \varphi'(\xi_n)(x_n - x^*) = \\ = \frac{p-1}{p}(x_n - x^*) + O(|x_n - x^*|^2)\,, \end{aligned}$$

где точка ξ_n лежит между x_n и x^*.

Опираясь на последнее равенство можно сделать следующие предположения:

– метод Ньютона не сходится, если $p \le 0.5$,

– метод Ньютона сходится, если $p > 0.5$, с линейной скоростью при $p \ne 1$ и $p \approx \dfrac{x_{n+1} - x_n}{2x_{n+1} - x_{n+2} - x_n}$,

– скорость сходимости метода Ньютона квадратичная при $p = 1$.

7.1. Метод Нютона с параметром

Рассмотрим частный случай уравнения (1) на интервале $[0,\,1]$

$$f(x) \equiv x^p = 0\,, \qquad p > 0\,,$$

т.е. $x^* = 0$ – корень кратности p функции $f(x)$. Метод Ньютона в этом случае

$$x_{n+1} = x_n - \frac{x_n^p}{p\,x_n^{p-1}} = \frac{p-1}{p}\,x_n$$

при $p \le 0.5$ не сходится, а при остальных $p \ne 1$ скорость его сходимости линейна.

Определим метод Ньютона с параметром:

$$x_{n+1} = x_n - p\,\frac{f(x_n)}{f'(x_n)}\,, \quad x_0 \in [\,a,\,b\,]\,, \quad n = 0,\,1,\,\ldots\,. \tag{6}$$

Очевидно, что для нашего примера этот метод дает решение за одну итерацию при любом $p > 0$.

Метод Ньютона с параметром можно трактовать как метод последовательных приближений решения уравнения

$$x = \psi(x) \equiv x - p\,\frac{f(x)}{f'(x)}\,. \tag{7}$$

Тогда, так как с учетом формул (5) имеем

$$\begin{aligned} &\psi(x) = (1-p)\,x + p\,\varphi(x)\,,\\ &\psi'(x) = (1-p) + p\left[\frac{p-1}{p} + O(|x-x^*|)\right] = O(|x_n - x^*|)\,, \end{aligned} \tag{8}$$

приближения x_n метода Ньютона с параметром к корню x^* функции $f(x)$ удовлетворяют соотношениям

$$\begin{aligned} x_{n+1} - x^* = \psi(x_n) - \psi(x^*) = \psi'(\xi_n)(x_n - x^*) =\\ = O(|x-x^*|)\,(x_n - x^*) = O(|x_n - x^*|^2)\,, \end{aligned}$$

где точка ξ_n лежит между x_n и x^*.

Опираясь на последнее равенство можно сделать следующее предположение:
– метод Ньютона с параметром сходится при любой положительной кратности p корня функции $f(x)$ и скорость его сходимости квадратичная. □

Если кратность p корня x^* функции $f(x)$ неизвестна, то разрешая (5) относительно p, получим

$$p = \frac{1}{1-\varphi'(x)} + O(|x - x^*|\,.$$

Тогда итерационный процесс

$$x_{n+1} = x_n - \frac{1}{1-\varphi'(x)} \frac{f(x_n)}{f'(x_n)}, \quad x_0 \in [\,a,\, b\,], \quad n = 0,\, 1,\, \ldots, \tag{9}$$

будет сходиться при достаточно близком к корню x^* приближении x_0, так как он является методом простой итерации для уравнения

$$x = \eta(x) \equiv \frac{\varphi(x) - x\,\varphi'(x)}{1-\varphi'(x)} \tag{10}$$

и $\eta'(x^*) = 0$. □

7.2. Две простые теоремы сходимости метода Ньютона

В заключение этого раздела докажем две простые теоремы, содержащие условия существования корня x^* функции $f(x)$ и сходимости к нему приближений метода Ньютона.

Теорема 1. *Если непрерывно дифференцируемая функция $f(x)$ выпукла на отрезке $[\,a,\, b\,]$, $f(a) > 0$ и $f(b) \le 0$, то уравнение*

$$f(x) = 0$$

имеет единственное решение $x^* \in [\,a,\, b\,]$, приближения x_n метода Ньютона

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \quad x_0 = a, \quad n = 0,\, 1,\, \ldots.$$

строго возрастают и сходятся к этому корню.

Доказательство. Существование решения уравнения $f(x) = 0$ следует из того, что непрерывная функция принимает все значения между $f(a) > 0$ и $f(b) \le 0$.

Пусть $x^* \le x^{**}$ – корни функции $f(x)$. Из определения выпуклости следует и условия $f(b) \le 0$ следует, что

$$0 = f(x^{**}) \le \frac{x^{**}-b}{x^*-b}\, f(x^*) - \frac{x^{**}-x^*}{b-x^*}\, f(b) = \frac{x^{**}-x^*}{b-x^*}\, f(b) \le 0,$$

и, значит, $x^* = x^{**}$, т.е. решения уравнения $f(x) = 0$ единственно.

Теперь покажем, что последовательность x_n метода Ньютона определена и строго возрастает. Прежде всего заметим, что на интервале $[\,a,\, x^*\,)$ функция $f(x) > 0$ и строго убывает. Первое следует из единственности корня, второе доказывается от противного: если $f(x+\varepsilon^2) \ge f(x)$ для некоторого $x+\varepsilon^2 \in [\,a,\, x^*\,)$, то из определения выпуклости и нашего предположения следует, что

$$\begin{aligned} f(x+\varepsilon^2) &\le \frac{x+\varepsilon^2-x^*}{x-x^*}\, f(x) + \frac{x+\varepsilon^2-x}{x^*-x}\, f(x^*) \\ &= \frac{(x^*-x)-\varepsilon^2}{x^*-x}\, f(x) < f(x) \le f(x+\varepsilon^2), \end{aligned}$$

т.е. получили противорчие.

Далее, $f'(x) < 0$ на интервале $[\,a,\, x^*\,)$. Действительно, так как

$$f(x+\varepsilon^2) \le \Big(1 - \frac{\varepsilon^2}{x^* - x}\Big) f(x)\,,$$

то

$$f'(x) = \lim_{\varepsilon^2 \to 0} \frac{f(x+\varepsilon^2) - f(x)}{\varepsilon^2} \le -\frac{f(x)}{x^* - x} < 0\,.$$

Отсюда следует, что

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} = x_n + \frac{|f(x_n)|}{|f'(x_n)|} \le x_n + |x^* - x_n| = x^*\,,$$

т.е. $x_{n+1} > x_n$ и $x_{n+1} \in [\,a,\, x^*\,]$.

Начиная с $x_0 = a$, мы получим ограниченную, строго возрастающую последовательность $\{x_n\}$ (конечную, если на некоторой итерации метода Ньютона $x_n = x^*$), предел которой существует и конечен, а переходя к пределу в левой и правой частях равенства

$$f'(x_n)\,(x_{n+1} - x_n) = f(x_n)\,,$$

убеждаемся, что он является корнем функции $f(x)$. □

Замечание. Аналогичные теоремы можно сформулировать и доказать для случая вогнутых функций и изменения функции от отрицательных значений к положительным. □

Пример. Пусть

$$P(t) = a_2\, t^2 - a_1\, t + a_0 \tag{11}$$

– полином второй степени с положительными корнями $t^* \le t^{**}$ и положительным коэффициентом a_2. Легко проверить, что на отрезке $[\,0,\, t^*\,]$ для этого полинома выполняются условия теоремы 1 и, следовательно, последовательность $\{t_n\}$ метода Ньютона

$$t_{n+1} = t_n - \frac{P(t_n)}{P'(t_n)}\,, \quad t_0 = 0\,, \quad n = 0,\, 1,\, \ldots\,, \tag{12}$$

строго возрастает и сходиться к t^*.

Если (12) переписать в виде

$$t_{n+1} = \varphi(t_n)\,, \quad t_0 = 0\,, \quad n = 0,\, 1,\, \ldots\,,$$

где

$$\varphi(t) = t - \frac{P(t)}{P'(t)} = t - \frac{(t - t^*)(t - t^{**})}{2\,t - (t^* + t^{**})}\,,$$

то

$$0 \le \varphi'(t) = \frac{2\,(t^* - t)\,(t^{**} - t)}{[(t^* - t) + (t^{**} - t)]^2} \le \frac{1}{2} \quad \forall\, t \in [\,0,\, t^*\,]\,,$$

и

$$0 \le t_{n+1} - t^* = \varphi(t_n) - \varphi(t^*) = \varphi'(\xi_n)\,(t_n - t^*) \le \frac{1}{2}\,(t_n - t^*)\,, \tag{13}$$

т.е. скорость сходимости метода Ньютона не хуже линейной (линейна при $t^* = t^{**}$.

Если $t^* < t^{**}$, то

$$0 \le \varphi'(t) = \frac{2\,(t^{**} - t)}{[(t^{**} - t) + (t^* - t)]^2}\,(t^* - t) \le \frac{2}{t^{**} - t}\,(t^* - t) \le$$
$$\le \frac{2}{t^{**} - t^*}\,(t^* - t) \quad \forall\, t \in [\,0,\, t^*\,]\,,$$

и

$$\begin{aligned} 0 \le t_{n+1} - t^* &= \varphi(t_n) - \varphi(t^*) = \varphi'(\xi_n)\,(t_n - t^*) \le \\ &\le \frac{2}{t^{**} - t^*}\,(t^* - \xi_n)\,(t_n - t^*) \le \frac{2}{t^{**} - t^*}\,(t_n - t^*)^2\,, \end{aligned} \tag{14}$$

т.е. скорость сходимости метода Ньютона квадратичная.

Заметим, что в этом примере выполняются условия :

$$\begin{aligned} a_2 &= \frac{1}{2}\,\max P''(t) > 0\,, \\ a_1 &= -P'(0) = \frac{t^* + t^{**}}{a_2} > 0\,, \\ a_0 &= P(0) = \frac{t^*\,t^{**}}{a_2} > 0\,. \end{aligned} \tag{15}$$

Теорема 2. *Пусть функция $f(x)$ определена и дважды непрерывно дифференцируема на отрезке $[\,x_0 - \varepsilon,\, x_0 + \varepsilon\,]$, $\varepsilon > 0$ и*

$$\begin{aligned} \frac{1}{2}\,\max_{|x - x_0| \le \varepsilon} |f''(x)| &\le a_2\,, \\ |f'(x_0)| &\ge a_1 > 0\,, \\ |f(x_0)| &\le a_0\,, \end{aligned}$$

а полином $P(t) = a_2\,t^2 - a_1\,t + a_0$ вида (11) имеет положительные корни $t^* \le t^{**}$ и $t^* \le \varepsilon$.

Тогда уравнение $f(x) = 0$ имеет на отрезке $[\,x_0 - \varepsilon,\, x_0 + \varepsilon\,]$ решение x^*, к которому сходится последовательность $\{x_n\}$ метода Ньютона

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

и имеет место оценка

$$|\,x_n - x^*\,| \le t^* - t_n \quad \forall\, n \le 0\,.$$

Доказательство. Прежде всего напомним, что последовательность

$$t_{n+1} = t_n - \frac{P(t_n)}{P'(t_n)}, \quad t_0 = 0, \quad n = 0, 1, \ldots,$$

строго возрастает и сходиться к t^*.

При $n = 0$ по условию теоремы $|f(x_0)| \le a_0$ и $|f'(x_0)| \ge a_1 > 0$, следовательно, x_1 определен и

$$|x_1 - x_0| = \Big|\frac{f(x_0)}{f'(x_0)}\Big| \le \frac{a_0}{a_1} = -\frac{P(0)}{P'(0)} = t_1 - t_0 < t^* \le \varepsilon\,.$$

Предположим, что для некоторого $n \ge 1$ доказано, что

$$\begin{aligned} |x_k - x_0| &\le \varepsilon\,, \\ |x_k - x_{k-1}| &\le t_k - t_{k-1}\,, \quad k = 1, \ldots n\,. \end{aligned} \tag{16}$$

Покажем, что эти неравенства будут выполняться при $k = n + 1$.

Для вычисления x_{n+1} необходимо и достаточно показать, что $f'(x_n) \ne 0$:

$$\begin{aligned} |f'(x_n)| &= |\,f'(x_0) + \int\limits_{x_0}^{x_n} f''(x)\,dx\,| \ \ge |\,f'(x_0)\,| - |\int\limits_{x_0}^{x_n} f''(x)\,dx\,| \ \ge \\ &\ge a_1 - 2\,a_2\,|x_n - x_0| \ge a_1 - 2\,a_2\,(|x_n - x_{n-1}| + \ldots + |x_1 - x_0|) \ge \\ &\ge a_1 - 2\,a_2(t_n - t_{n-1} + \ldots + t_1 - t_0) = a_1 - 2\,a_2\,t_n = \\ &= -P'(t_n) > 0\,. \end{aligned}$$

Теперь оценим

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

Так как $|f'(x_n)| \ge -P'(t_n)$ и

$$\begin{aligned} |\,f(x_n)\,| &= |\,f(x_{n-1}) - f'(x_{n-1})\,(x_n - x_{n-1}) + \frac{f''(x_{n-1})}{2}\,(x_n - x_{n-1})^2\,| = \\ &= |\,\frac{f''(x_{n-1})}{2}\,(x_n - x_{n-1})^2\,| \le a_2\,(t_n - t_{n-1})^2\,, \\ P(t_n) &= P(t_{n-1}) - P'(t_{n-1})\,(t_n - t_{n-1}) + \frac{P''(t_{n-1})}{2}\,(t_n - t_{n-1})^2 = \\ &= \frac{P''(t_{n-1})}{2}\,(t_n - t_{n-1})^2 \,=\, a_2\,(t_n - t_{n-1})^2\,, \end{aligned}$$

то

$$|\,x_{n+1} - x_n\,| \,=\, |\,\frac{f(x_n)}{f'(x_n)}\,| \,\le\, \frac{P(t_n)}{-P'(t_n)} \,=\, t_{n+1} - t_n\,,$$

т.е. выполняется второе из предположений (16).

Оценим $|x_{n+1} - x_n|$:

$$
\begin{aligned}
|\, x_{n+1} - x_n \,| \le |\, x_{n+1} - x_{n-1} + \ldots + |\, x_1 - x_0 \,| \le \\
\le t_{n+1} - t_{n-1} + \ldots + t_1 - t_0 = t_{n+1} \le t^* \le \varepsilon \,.
\end{aligned}
$$

Следовательно, неравенства (16) выполняются при любом $n \ge 0$.

Последовательность $\{x_n\}$ является фундаментальной по Коши, так как

$$
\begin{aligned}
|\, x_{n+m} - x_n \,| \le |\, x_{n+m} - x_{n+m-1} + \ldots + |\, x_{n+1} - x_n \,| \le \\
\le t_{n+m} - t_{n+m-1} + \ldots + t_{n+1} - t_n = t_{n+m} - t_n \,,
\end{aligned}
\tag{17}
$$

а последовательность $\{t_n\}$ фундаментальна, так как она сходится. Следовательно, $x_n \to x^* \in [\, x_0 - \varepsilon, \, x_0 + \varepsilon \,]$, а так как $f(x)$ и $f'(x)$ непрерывны, то, переходя к пределу при $n \to \infty$ в равенстве

$$
f'(x_n)\,(x_{n+1} - x_n) \;=\; f(x_n)\,,
$$

получим $f(x^*) = 0$. Переходя в неравенствах (17) к пределу при $m \to \infty$, получим последнее утверждение теоремы. □

8. Численные методы решения задачи Коши для обыкновенных дифференциальных уравнений

Результатом математического моделирования реальных физических процессов во многих случаях являются одно или несколько (система) обыкновенных дифференциальных уравнений. Мы ограничимся изучением численных методов решения задачи Коши для одного уравнения :

$$\begin{aligned} \frac{d\,u(t)}{d\,t} &= f(t,\,u(t)\,)\,, \quad t > 0\,, \\ u(0) &= u_0\,, \end{aligned} \tag{1}$$

предполагая, что правая часть $f(t,\,u)$ дифференциального уравнения является известной, достаточно гладкой функцией двух переменных t и u, а решение задачи $u(t)$ задано в начальный момент времени.

Напомним методику применения принципа сжимающего отображения для доказательства существования решения задачи Коши (1).

Теорема 1. *Пусть функция $f(t,u)$ определена и непрерывна при $t \in [\,0,\,T\,]$ и $u \in [\,u_0 - a,\,u_0 + a\,]$, непрерывна по Липшицу по второму аргументу с постоянной L, независящей от t :*

$$|\,f(t,\,u) - f(t,\,v) \le L\,|\,u - v\,| \quad \forall\,u,\,v \in [\,u_0 - a,\,u_0 + a\,] \quad \forall\,t \in [\,0,\,T\,]\,,$$

тогда задача Коши (1) на интервале $[\,0,\,t_0]$, где $t_0 < \min\left\{\,T,\,\frac{a}{M},\,\frac{1}{L}\,\right\}$, M – максимум модуля функции $f(t,u)$ в области ее задания, имеет единственное решение.

Доказательство. Проинтегрируем дифференциальное уравнение из (1) :

$$u(t) = u_0 + \int\limits_0^t f(t,\,u(t)\,)\,dt\,.$$

Правую часть этого выражения мы можем рассматривать как отображение (преобразование) $\varphi(u)$ непрерывной функции $u(t)$ как элемента полного метрического пространства $\mathcal{M}$ всех непрерывных на отрезке $[\,0,\,t_0\,]$ функций, значения которых принадлежат интервалу $[\,u_0 - a,\,u_0 + a\,]$, с расстоянием $\rho(u,v) = \max\limits_{0 \le t \le t_0} |u(t) - v(t)|$, в непрерывную функцию.

Выберем t_0 так, чтобы $\varphi(u)$ отображало $\mathcal{M}$ в $\mathcal{M}$ и было оператором сжатия. Для выполнения первого условия достаточно потребовать, чтобы

$$|\,\varphi(u) - u_0\,| = |\int\limits_0^t f(t,\,u(t)\,)\,dt\,| \le t_0 \max_{0 \le t \le t_0} |f(t,u(t))| \le t_0\,M \le a\,.$$

Для выполнения второго условия достаточно потребовать, чтобы

$$| \varphi(u) - \varphi(v) | = | \int\limits_0^t [f(t, u(t)) - f(t, v(t))] \, dt | \le$$

$$\le \int\limits_0^t L \, | u(t) - v(t) | \, dt \le t_0 \, L \, \rho(u,v) < \rho(u,v) \, .$$

Очевидно, что при любом $t_0 < \min \left\{ T, \frac{a}{M}, \frac{1}{L} \right\}$ $\varphi(u)$ – сжимающее отображение из $\mathcal{M}$ в $\mathcal{M}$ и, следовательно, имеет единственную неподвижную точку $u(t) \in \mathcal{M}$. □

Из этой теоремы (вернее из принципа сжимающих отображений) следует сходимость метода последовательных приближений

$$u_{n+1}(t) = u_0 + \int\limits_0^t f(t, u_n(t)) \, dt \, , \quad u_0(t) \in \mathcal{M} \, , \quad n = 0, 1, \ldots ,$$

практическое применение которого сталкивается с усложняющейся от шага к шагу задачей интегрирования.

Заметим, что при $t = 0$ мы можем вычислить "все"производные искомого решения $u(t)$ задачи Коши (1):

$$\begin{aligned} u(0) &= u_0 \, , \\ u'(0) &= f(t, u(t)) \big|_{t=0} \, , \\ u''(0) &= \Big[\frac{\partial f(t,u)}{\partial t} + \frac{\partial f(t,u)}{\partial u} \, u'(t) \Big] \Big|_{t=0} \, , \\ &\cdot \ \cdot \ \cdot \ \cdot \ \cdot \ \cdot \ \cdot \ , \end{aligned}$$

и, стало быть, при малых значениях t разложение в ряд Тейлора $u(t) = u(0) + u'(0)\, t + 0.5\, u''(0)\, t^2 + \ldots$ может использоваться для приближения решения задачи Коши (1).

8.1. Методы Рунге – Кутта

Идея построения методов Рунге – Кутта приближенного решения задачи Коши:

$$\begin{aligned} \frac{d\,u(t)}{d\,t} &= f(t, u(t)) \, , \quad t > 0 \, , \\ u(0) &= u_0 \, , \end{aligned} \tag{1}$$

заключается в следующем.

Пусть $t_0 = 0$, $t_1 = \tau$, ..., $t_n = n\,\tau$, ... – равномерная сетка по переменной t с шагом (интегрирования) τ. Предположим, что нам известно точное или приближенное значение $y_n = y_n(t_n)$ решения задачи (1) при $t = t_n$. Будем считать, что на интервале $[\, t_n, t_n + \tau \,]$ функция $y_n(t)$ удовлетворяет дифференциальному уравнению из (1) и, следовательно,

$$y_n(t_n+\tau)=y_n+\int\limits_{t_n}^{t_n+\tau} f(t,\,y_n(t)\,)\,dt\,. \tag{2}$$

Будем считать заданными m точек на интервале $[\,t_n,\,t_n+\tau\,]$:

$$t_n\,\le\,t_n+\alpha_1\,\tau\,<\,t_n+\alpha_2\,\tau\,<\,\ldots\,<\,t_n+\alpha_m\,\tau\le t_n+\tau\,.$$

Интеграл в правой части равенства (2) заменим квадратурной формулой

$$\int\limits_{t_n}^{t_n+\tau} f(t,\,y_n(t)\,)\,dt\approx\tau\sum_{i=1}^{m}\sigma_i f(t_n+\alpha_i\,\tau,\,y_n(t_n+\alpha_i\,\tau)\,)\,,$$

а для определения неизвестных $y_n(t_n+\alpha_i\,\tau)$ в равенствах

$$y_n(t_n+\alpha_i\,\tau)=y_n+\int\limits_{t_n}^{t_n+\alpha_i\,\tau} f(t,\,y_n(t)\,)\,dt$$

заменим интегралы квадратурными формулами

$$\int\limits_{t_n}^{t_n+\alpha_i\,\tau} f(t,\,y_n(t)\,)\,dt\approx\tau\sum_{j=1}^{m}\sigma_{ij} f(t_n+\alpha_j\,\tau,\,y_n(t_n+\alpha_j\,\tau)\,)\,,$$
$$i=1,\,2,\,\ldots,\,m\,,$$

что приведет нас к системе относительно $\tilde y_i\approx y_n(t_n+\alpha_i\,\tau)$:

$$\tilde y_i=y_n+\tau\sum_{j=1}^{m}\sigma_{ij} f(t_n+\alpha_j\,\tau,\,\tilde y_j\,)\,,$$
$$i=1,\,2,\,\ldots,\,m\,,$$

решив которую, получим приближение

$$y_{n+1}=y_n+\tau\sum_{i=1}^{m}\sigma_i f(t_n+\alpha_i\,\tau,\,\tilde y_i\,)\approx y_n(t_n+\tau)$$

к решению задачи Коши (1) при $t=t_{n+1}$.

Введем обозначение $k_i=k_i(y_n)=f(t_n+\alpha_i\,\tau,\,\tilde y_i\,)$, тогда

$$k_i=f(t_n+\alpha_i\,\tau,\,y_n+\tau\sum_{j=1}^{m}\sigma_{ij}k_j\,)\,,\qquad i=1,\,2,\,\ldots,\,m\,, \tag{3}$$

$$y_{n+1} = y_n + \tau \sum_{i=1}^{m} \sigma_i \, k_i \tag{4}$$

– m-этапный метод Рунге – Кутта с параметрами $\{\alpha_i\}_{i=1}^m$, $\{\sigma_{ij}\}_{i,j=1}^m$, $\{\sigma_i\}_{i=1}^m$.

Метод Рунге – Кутта называется явным, если матрица его параметров $(\sigma_{ij})_{i,j=1}^m$ является строго нижней треугольной матрицей, т.е. значения k_1, ..., k_m явно вычисляются одно за другим по формулам (3) (обычно в явных методах параметр α_1 также выбирается равным нулю); в противном случае метод называется неявным. □

Прежде всего необходимо установить разрешимость системы (3) и указать метод ее решения.

Теорема 2. *Если функция $f(t,\, u(t)\,)$ непрерывна по Липшицу, то существует $\tau_0 > 0$ такое, что при любом $\tau \in (\,0,\, \tau_0\,]$ система (3) разрешима, а метод последовательных приближений сходится к ее решению.*

Доказательство. Для большей наглядности ограничимся случаем $m = 2$:

$$k_1 = f(t_n + \alpha_1 \tau,\; y_n + \tau\, \sigma_{11} k_1 + \tau\, \sigma_{12} k_2\,)\,,$$
$$k_2 = f(t_n + \alpha_2 \tau,\; y_n + \tau\, \sigma_{21} k_1 + \tau\, \sigma_{22} k_2\,)$$

или

$$k \equiv \begin{bmatrix} k1 \\ k2 \end{bmatrix} = \begin{bmatrix} f(t_n + \alpha_1 \tau,\; y_n + \tau\, \sigma_{11} k_1 + \tau\, \sigma_{12} k_2\,) \\ f(t_n + \alpha_2 \tau,\; y_n + \tau\, \sigma_{21} k_1 + \tau\, \sigma_{22} k_2\,) \end{bmatrix} \equiv F(k) \tag{5}$$

Метод последовательных приближений

$$k^{n+1} = F(k^n)\,, \quad k^0 = \begin{bmatrix} f(t_n + \alpha_1 \tau,\; y_n\,) \\ f(t_n + \alpha_2 \tau,\; y_n\,) \end{bmatrix}\,, \quad n = 0,\, 1,\, \ldots, \tag{6}$$

будет сходится к решению (единственному) уравнения (5), если $F(k)$ будет сжимающим отображением из некоторого полного метрического пространства $\mathcal{M}$ в $\mathcal{M}$. Пусть d – произвольное положительное число и

$$\mathcal{M} \;= \{\, k \in R^2 : \quad \rho(k,\, k^0) = ||\, k - k^0\, ||_\infty \le d\,\}\,.$$

Нетрудно показать, что $\mathcal{M}$ является полным метрическим пространством.

Предполагая, что функция $f(t,y)$ непрерывна по Липшицу по второму аргументу с постоянной L, сначала найдем условие, при котором $F(k) \in \mathcal{M}$ для всех $k \in \mathcal{M}$:

$$\begin{aligned}
&||\, F(k) - k^0\, ||_\infty = \\
&= \max\{|\, f(t_n + \alpha_1 \tau,\; y_n + \tau\, \sigma_{11} k_1 + \tau\, \sigma_{12} k_2\,) - f(t_n + \alpha_1 \tau,\; y_n\,)\,|\,, \\
&\qquad\quad |\, f(t_n + \alpha_2 \tau,\; y_n + \tau\, \sigma_{21} k_1 + \tau\, \sigma_{22} k_2\,) - f(t_n + \alpha_2 \tau,\; y_n\,)\,|\,\} \le \\
&\le \max\{L\,|\,\tau\, \sigma_{11} k_1 + \tau\, \sigma_{12} k_2\,|\,, L\,|\,\tau\, \sigma_{21} k_1 + \tau\, \sigma_{22} k_2\,|\,\} \le \\
&\qquad\quad \le L\,\tau\, \max\{|\,\sigma_{11}\,| + |\,\sigma_{12}\,|\,, |\,\sigma_{21}\,| + |\,\sigma_{22}\,|\,\}\, ||\, k\, ||_\infty\,.
\end{aligned}$$

Так как $||\, k - k^0\, ||_\infty \le d$, то

$$
\begin{aligned}
||\, F(k) - k^0 \,||_\infty &\le L\,\tau\,\sigma\,(\,d + ||\,k^0\,||_\infty\,)\,,\\
\sigma \ &= \max\{|\,\sigma_{11}\,| + |\,\sigma_{12}\,|\,, |\,\sigma_{21}\,| + |\,\sigma_{22}\,|\,\}\,,
\end{aligned}
$$

и, выбирая τ достаточно малым, получим нужное свойство отображения $F(k)$.

Теперь покажем, что при достаточно малых τ отображение $F(k)$ будет сжимающим. Пусть $k, l \in \mathcal{M}$, тогда

$$
\begin{aligned}
&||\, F(k) - F(l)\,||_\infty = \max\{\\
&|f(t_n + \alpha_1\tau, y_n + \tau\sigma_{11}k_1 + \tau\sigma_{12}k_2) - f(t_n + \alpha_1\tau, y_n + \tau\sigma_{11}l_1 + \tau\sigma_{12}l_2)|\,,\\
&|f(t_n + \alpha_2\tau, y_n + \tau\sigma_{21}k_1 + \tau\sigma_{22}k_2) - f(t_n + \alpha_2\tau, y_n + \tau\sigma_{21}l_1 + \tau\sigma_{22}l_2)|\}\\
&\qquad\le L\,\tau\,\max\{|\,\sigma_{11}\,| + |\,\sigma_{12}\,|\,, |\,\sigma_{21}\,| + |\,\sigma_{22}\,|\,\}\,||\,k - l\,||_\infty
\end{aligned}
$$

и, выбирая $\tau < (L\,\sigma)^{-1}$, получим условие сжатия отображения $F(k)$. □

Изучим поведение погрешности решения $z_n \equiv y_n - u(t_n)$. Обозначим через $k_i^* \equiv k_i(u(t_n))$ решение системы

$$
k_i^* = f(t_n + \alpha_i\,\tau,\, u(t_n) + \tau\sum_{j=1}^{m}\sigma_{ij}k_j^*\,)\,,
$$
$$
i = 1,\, 2,\, \dots,\, m\,.
$$

Легко получить формулу для пересчета погрешности решения:

$$
z_{n+1} = z_n + \tau\sum_{i=1}^{m}\sigma_i(k_i - k_i^*) - \tau\,\psi_n\,, \tag{7}
$$

где

$$
\psi_n = \frac{u(t_{n+1}) - u(t_n)}{\tau} - \sum_{i=1}^{m}\sigma_i k_i(u(t_n)) \tag{8}
$$

– погрешность аппроксимации уравнения (1) на его решении.

Теорема 3. *Если на конечном интервале $[\,0,\, T\,]$ задача Коши (1) имеет решение, а функция $f(t,\, u(t)\,)$ непрерывна по Липшицу, то существует $\tau_0 > 0$ такое, что при любом $\tau \in (\,0,\, \tau_0\,]$ для погрешности ее решения методом Рунге – Кутта справедлива оценка*

$$
\begin{gathered}
|z_n| \le \mathrm{e}^{CT}\big(\,|z_0| + T\max_{1\le i\le n}|\psi_i|\,\big) \le \mathrm{e}^{CT}\big(\,|z_0| + T\,||\psi_i||_\infty\,\big)\\
\forall\, t_n \in [\,0,\, T\,]
\end{gathered} \tag{9}
$$

с некоторой постоянной C.

Доказательство. Оценим разность $(k_i - k_i^*)$ из (7):

$$
\begin{aligned}
&|\,k_i - k_i^*\,| = \\
&= |f(t_n + \alpha_i\tau,\, y_n + \tau \sum_{j=1}^{m} \sigma_{ij} k_j) - f(t_n + \alpha_i\tau,\, u(t_n) + \tau \sum_{j=1}^{m} \sigma_{ij} k_j^*)| \le \\
&\qquad \le L\,|z_n| + L\,\tau \sum_{j=1}^{m} |\sigma_{ij}|\,|k_j - k_j^*| \le \\
&\qquad \le L\,|z_n| + \tau\Big(L \sum_{j=1}^{m} |\sigma_{ij}|\Big)\,||k - k^*||_\infty\,, \\
&\qquad\qquad\qquad i \,=\, 1,\, \dots\, m\,.
\end{aligned}
$$

Из этих неравенств следует, что, во-первых:

$$
||\,k - k^*\,||_\infty \le \frac{L}{1 - \tau\,L \max\limits_{1\le i\le m} \sum\limits_{j=1}^{m} |\sigma_{ij}|}\,|z_n|
$$

и, во-вторых:

$$
|\sum_{i=1}^{m} \sigma_i (k_i - k_i^*)| \le \frac{L \sum\limits_{i=1}^{m} |\sigma_i|}{1 - \tau\,L \max\limits_{1\le i\le m} \sum\limits_{j=1}^{m} |\sigma_{ij}|}\,|z_n| \le C\,|z_n|
$$

для всех $\tau \in (\,0,\, \tau_0]$, где τ_0 определяется из условия положительности знаменателя этих неравенств.

Используя последнее неравенство, из (7) получим, что

$$
\begin{aligned}
|z_{n+1}| &\le |z_n| + \tau\,C\,|z_n| + \tau\,|\psi_n| = (1 + \tau\,C)\,|z_n| + \tau\,|\psi_n| \le \\
&\le (1 + \tau\,C)^2\,|z_{n-1}| + (1 + \tau\,C)\,\tau\,|\psi_{n-1}| + \tau\,|\psi_n| \le \\
&\quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
&\le (1 + \tau\,C)^{n+1}\,|z_0| + \tau \sum_{i=1}^{n} (1 + \tau\,C)^{n-i}\,|\psi_i|\,.
\end{aligned} \tag{10}
$$

Тогда, так как для любого положительного $x = C\tau$ имеет место неравенство $(1 + x)^{\frac{1}{x}} \le \mathrm{e}$, то на любом конечном интервале $[\,0,\, T\,]$ только что полученное неравенство можно переписать в следующем виде:

$$
|z_n| \le \mathrm{e}^{CT}\big(\,|z_0| + T \max_{1\le i\le n} |\psi_i|\,\big) \le \mathrm{e}^{CT}\big(\,|z_0| + T\,||\psi_i||_\infty\,\big)\,,
$$

т.е. погрешность решения совпадает по порядку малости с погрешностью приближения начального условия и максимальной ошибкой аппроксимации дифференциального уравнения на его решении. □

8.2. Построение методов Рунге – Кутта

Из теоремы 3 следует, что при решении задачи Коши

$$
\begin{aligned}
&u'(t) = f(t,\, u(t)\,)\,, \quad t > 0\,,\\
&u(0) = u_0
\end{aligned}
\tag{11}
$$

m-этапным методом Рунге – Кутта

$$
\begin{aligned}
&k_i(\tau, y_n) = f(t_n + \alpha_i\,\tau,\, y_n + \tau \sum_{j=1}^{m} \sigma_{ij} k_j(\tau, y_n)\,)\,, \quad i = 1,\, 2,\, \dots,\, m\,,\\
&y_0 = u_0\,, \quad y_{n+1} = y_n + \tau \sum_{i=1}^{m} \sigma_i\, k_i(\tau, y_n)\,, \quad n = 0,\, 1,\, \dots,
\end{aligned}
\tag{12}
$$

параметры метода

$$
\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{bmatrix}, \quad
\begin{bmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1m} \\ \sigma_{21} & \sigma_{22} & \dots & \sigma_{2m} \\ \dots & \dots & \dots & \dots \\ \sigma_{m1} & \sigma_{m2} & \dots & \sigma_{mm} \end{bmatrix}, \quad
\begin{bmatrix} \sigma_1 \\ \sigma_2 \\ \vdots \\ \sigma_m \end{bmatrix}
\tag{13}
$$

следует выбирать из условия максимальности порядка малости по τ погрешности аппроксимации

$$
\psi_n(\tau) = \frac{u(t_n + \tau) - u(t_n)}{\tau} - \sum_{i=1}^{m} \sigma_i k_i(\tau, u(t_n))
\tag{14}
$$

на решении задачи (11).

Очевидно, что, если функция $f(t,u)$ p раз непрерывно дифференцируема, то $u(t)$ (так как $u'(t) = f(t,\, u(t)\,)$) непрерывно дифференцируема $p+1$ раз и

$$
\begin{aligned}
\psi_n(\tau) = &\Big[\, u'(t_n) - \sum_{i=1}^{m} \sigma_i k_i(0, u(t_n)\,) \Big] + \\
&+ \Big[\, \frac{1}{2}\, u''(t_n) - \sum_{i=1}^{m} \sigma_i k_i'(0, u(t_n)\,) \Big]\, \tau \;+\; \dots + \\
&+ \Big[\, \frac{1}{p}\, u^{(p)}(t_n) - \sum_{i=1}^{m} \sigma_i k_i^{(p-1)}(0, u(t_n)\,) \Big]\, \frac{\tau^{p-1}}{(p-1)!} + O(\tau^p)
\end{aligned}
\tag{15}
$$

и, следовательно, если мы хотим построить метод p-того порядка точности, нужно найти параметры (13) такие, чтобы выражения в квадратных скобках разложения (15) обратились в нуль для любой функции $f(t,u)$.

8.3. Метод Эйлера

Пусть $m = 1$, $\alpha_1 = 0$, $\sigma_{11} = 0$ и $\sigma_1 = 1$, тогда метод Эйлера:

$$
y_0 = u_0\,, \quad y_{n+1} = y_n + \tau\, f(t_n, y_n)\,, \quad n = 0,\, 1,\, \dots,
\tag{16}
$$

имеет первый порядок точности по τ.

Действительно, выражение (15) для погрешности аппроксимации может быть представлено в виде

$$\psi_n(\tau) = \Big[\, u'(t_n) - f(t_n, u(t_n)\,)\Big] \;+\; \Big[\,\frac{1}{2}\,u''(t_n)\Big]\,\tau \;+\; O(\tau^2) \;=$$
$$= \frac{1}{2}\,\Big[\,\frac{\partial f(t_n, u(t_n)\,)}{\partial t} \;+\; \frac{\partial f(t_n, u(t_n)\,)}{\partial u}\, f(t_n, u(t_n)\,)\Big]\,\tau \;+\; O(\tau^2)$$

и, следовательно теорема 8.3 гарантирует первый порядок точности по τ метода (16).

8.4. Семейство методов второго порядка точности

Пусть $m = 2$, тогда метод Рунге – Кутта:

$$\begin{aligned} & k_1(\tau, y_n) = f(t_n + \alpha_1\tau, y_n + \tau(\sigma_{11}k_1 + \sigma_{12}k_2))\,, \\ & k_2(\tau, y_n) = f(t_n + \alpha_2\tau, y_n + \tau(\sigma_{21}k_1 + \sigma_{22}k_2))\,, \\ y_0 = u_0\,, \quad & y_{n+1} = y_n + \tau\,(\,\sigma_1 k_1(\tau, y_n) + \sigma_2 k_2(\tau, y_n)\,)\,, \quad n = 0,\, 1,\, \ldots\,, \end{aligned} \tag{17}$$

будет иметь второй порядок точности по τ, если два первых слагаемых в разложении (15) для погрешности аппроксимации равны нулю:

$$\begin{aligned} u'(t_n) - \sigma_1 k_1(0, u(t_n)\,) - \sigma_2 k_2(0, u(t_n)\,) = 0 \\ \frac{1}{2}\,u''(t_n) - \sigma_1 k_1'(0, u(t_n)\,) - \sigma_2 k_2'(0, u(t_n)\,) = 0\,. \end{aligned} \tag{18}$$

Так как

$$\begin{aligned} u'(t_n) &= f(t_n, u(t_n)\,)\,, \\ u''(t_n) &= f_t(t_n, u(t_n)\,) + f_u(t_n, u(t_n)\,)\, f(t_n, u(t_n)\,)\,, \\ k_1(0, u(t_n)\,) &= f(t_n, u(t_n)\,)\,, \\ k_2(0, u(t_n)\,) &= f(t_n, u(t_n)\,)\,, \\ k_1'(0, u(t_n)\,) &= f_t(t_n, u(t_n)\,)\,\alpha_1 + f_u(t_n, u(t_n)\,)\,[\,\sigma_{11}k_1 + \sigma_{12}k_2\,]\,, \\ k_2'(0, u(t_n)\,) &= f_t(t_n, u(t_n)\,)\,\alpha_2 + f_u(t_n, u(t_n)\,)\,[\,\sigma_{21}k_1 + \sigma_{22}k_2\,]\,, \end{aligned}$$

то, опуская аргументы у функции f и ее частных производных, условия (18) можно переписать в виде

$$\begin{aligned} [\,1 - \sigma_1 - \sigma_2\,]\, f = 0\,, \\ \Big\{\,[\,0.5 - \sigma_1\alpha_1 - \sigma_2\alpha_2\,]\, f_t + \\ + [\,0.5 - \sigma_1(\sigma_{11} + \sigma_{12}) - \sigma_2(\sigma_{21} + \sigma_{22})\,]\,\Big\}\, f_u\, f = 0\,. \end{aligned}$$

Отсюда следует, что метод Рунге – Кутта (17) будет иметь второй порядок точности по τ при любой гладкой правой части $f(t, u)$ задачи Коши (11), если его параметры удовлетворяют системе уравнений

$$\begin{aligned} \sigma_1 \;+\; \sigma_2 = 1\,, \\ \sigma_1\,\alpha_1 \;+\; \sigma_2\,\alpha_2 = 0.5\,, \\ \sigma_1\,(\sigma_{11} + \sigma_{12}) \;+\; \sigma_2\,(\sigma_{21} + \sigma_{22}) = 0.5\,. \end{aligned}$$

Частное (однопараметрическое) решение этой системы:

$$\sigma_1 = 1 - \sigma\,, \quad \sigma_2 = \sigma\,,$$
$$\alpha_1 = \sigma_{11} = \sigma_{12} = \sigma_{22} = 0\,,$$
$$\alpha_2 = 0.5\sigma^{-1}\,, \quad \sigma_{21} = 0.5\sigma^{-1}$$

определяет семейство явных 2-х этапных методов Рунге – Кутта

$$y_{n+1} = y_n + \tau\left[\,(1-\sigma)\,f(t_n, y_n) + \sigma\, f\big(t_n + \tfrac{\tau}{2\sigma},\, y_n + \tfrac{\tau}{2\sigma}\, f(t_n, y_n)\big)\,\right],$$
$$y_0 = u_0\,, \quad n = 0,\, 1,\, \ldots .$$

второго порядка точности при любом фиксированном $\sigma \neq 0$. Заметим, что при $\sigma \to 0$ этот метод превращается в метод Эйлера первого порядка точности. $\square$

9. Многошаговые разностные методы решения задачи Коши для ОДУ

Трудоемкость численных методов решения задачи Коши для обыкновенных дифференциальных уравнений (ОДУ) чаще всего определяется количеством вычислений правой части на каждом шаге метода. Для явного m-этапного метода Рунге – Кутта эта характеристика равна m и анализ формул метода показывает, что вычисленные значения правой части на одном шаге не используются на другом. Построение методов, использующих ранее вычисленные значения правой части, позволяет уменьшить трудоемкость приближенного решения задачи Коши для ОДУ.

Линейным m-шаговым разностным методом решения задачи Коши

$$\begin{aligned} \frac{d\,u(t)}{d\,t} &= f(t,\, u(t)\,)\,, \quad t > 0\,, \\ u(0) &= u_0\,, \end{aligned} \tag{1}$$

называется система разностных уравнений

$$\begin{gathered} \frac{a_0 y_n + a_1 y_{n-1} + \ldots + a_m y_{n-m}}{\tau} = b_0 f_n + b_1 f_{n-1} + \ldots + b_m f_{n-m}\,, \\ n = m,\, m+1,\, \ldots\,, \end{gathered} \tag{2}$$

где $a_0 \neq 0$, $a_1, \ldots, a_m$ и $b_0, b_1, \ldots, b_m$ – числовые параметры метода, не зависящие от n; τ – шаг сетки $\{t_n = n \cdot \tau\}$; $f_n = f(t_n, y_n)$ – значения правой части; y_n – приближения к значениям решения $u(t_n)$; единственное решение которой (по крайней мере в случае $b_0 = 0$), определяется по начальным данным $y_0, y_1, \ldots, y_{m-1}$, которые можно вычислить, например, методом Рунге – Кутта. □

Линейный m-шаговый разностный метод (2) называется явным, если $b_0 = 0$. Очевидно, что его трудоемкость равна единице. □

При $b_0 \neq 0$ метод (2) называется неявным, поскольку для определения $y = y_n$ приходится решать нелинейное уравнение

$$y - \tau\,\frac{b_0}{a_0}\,f(t_n, y) \;=\; \sum_{k=1}^{m} \Big(\,\tau\,\frac{b_k}{a_0}\,f_{n-k} - \frac{a_k}{a_0}\,y_{n-k}\Big)\,, \tag{3}$$

например, методом Ньютона, выбирая $y^{(0)} = y_{n-1}$. □

9.1. Сходимость многошаговых разностных методов

Погрешность решения $z_n = y_n - u(t_n)$ задачи Коши (1) m-шаговым разностным методом (2), как и в случае методов Рунге – Кутта, по порядку малости шага интегрирования τ не превышает ошибок вычисления начальных данных и аппроксимации дифференциального уравнения

$$\psi_n = \frac{a_0 u(t_n) + a_1 u(t_{n-1}) + \ldots + a_m u(t_{n-m})}{\tau} -$$
$$- \left[\, b_0 f(t_n, u(t_n)) + b_1 f(t_{n-1}, u(t_{n-1})) + \ldots + b_m f(t_{n-m}, u(t_{n-m})) \,\right].$$

Легко проверить, что

$$z_n = -\sum_{k=1}^{m} \frac{a_k}{a_0} z_{n-k} + \frac{\tau}{a_0} (\varphi_n + \psi_n)\,, \tag{4}$$

где

$$\varphi_n = \sum_{k=0}^{m} b_k \left[\, f(t_{n-k}, y_{n-k}) - f(t_{n-k}, u(_{n-k})) \,\right],$$

а ψ_n – погрешность аппроксимации.

Дополним (4) тождествами $z_{n-1} = z_{n-1}, \ldots, z_{n-m+1} = z_{n-m+1}$ и перепишем в матричном виде

$$Z^n \,=\, S\, Z^{n-1} \,+\, \frac{\tau}{a_0} \left(\Phi^n \,+\, \Psi^n \right), \tag{5}$$

где

$$Z^{n-1} = \begin{bmatrix} z_{n-1} \\ z_{n-2} \\ \vdots \\ z_{n-m} \end{bmatrix}, \quad \Phi^n = \begin{bmatrix} \varphi_n \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \Psi^n = \begin{bmatrix} \psi_n \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

$$S = \begin{bmatrix} p_1 & p_2 & \ldots & p_{m-1} & p_m \\ 1 & 0 & \ldots & 0 & 0 \\ 0 & 1 & \ldots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \ldots & 1 & 0 \end{bmatrix}, \quad p_i = -\frac{a_i}{a_0}\,.$$

Формулу (5) (одношаговую) принято называть каноническим видом m-шагового разностного метода (2) для погрешности решения с начальным вектором Z^{m-1} размерности m, а матрицу S порядка m – оператором шага.

Характеристический полином матрицы S с точностью до множителя a_0 совпадает с полиномом

$$P_m(x) \,=\, a_0\, x^m \,+\, a_1\, x^{m-1} \,+\, \ldots \,+\, a_m\,, \tag{6}$$

который называется характеристическим полиномом m - шагового разностного метода (2).

Если у этого полинома есть корень λ (собственное значение матрицы S), модуль которого больше единицы, то ему соответствует собственный вектор Z_λ и, если случайно окажется, что $Z^{m-1} = Z_\lambda$, в погрешности решения Z^{n+m-1} будет присутствовать быстро растущее (по норме) слагаемое $\lambda^n Z_\lambda$. Поскольку начальный вектор Z^{m-1} практически всегда будет содержать в своем разложении по собственным векторам матрицы S составляющую по Z_λ (или она появится за счет ошибок округления), то погрешность Z^n будет быстро (по норме) расти.

Следовательно, для устойчивости вычислений необходимо, чтобы все корни полинома (6) не превосходили по модулю единицы.

Этого условия недостаточно. Если у полинома (6) есть корень λ, модуль которого равен единице, кратности большей единицы, то ему соответствует только один собственный вектор Z_λ, так как ранг матрицы

$$S - \lambda\, E = \left[\begin{array}{ccccc} p_1 - \lambda & p_2 & \dots & p_{m-1} & p_m \\ 1 & -\lambda & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \multicolumn{5}{c}{\dotfill} \\ 0 & 0 & \dots & 1 & -\lambda \end{array} \right]$$

равен $m-1$ (последние $m-1$ строки линейно независимы).

Следовательно, в жордановой форме $\tilde{S} = Q\, S\, Q^{-1}$ будет присутствовать жорданов блок вида (кратность корня равна двум)

$$\left[\begin{array}{cc} \lambda & 1 \\ 0 & \lambda \end{array} \right], \quad \left[\begin{array}{cc} \lambda & 1 \\ 0 & \lambda \end{array} \right]^n = \left[\begin{array}{cc} \lambda^n & n\lambda^{n-1} \\ 0 & \lambda^n \end{array} \right],$$

что приведет к росту (по норме) входящего в погрешность решения Z^{n+m-1} слагаемого $S^n\, Z^{m-1}$.

Следовательно, для устойчивости вычислений необходимо, чтобы корни полинома (6), равные по модулю единице, были простыми.

Определение 1. Многошаговый разностный метод (2) удовлетворяет условию корней, если все корни его характеристического полинома (6) по модулю не превосходят единицу, а равные по модулю единице – простые.

Лемма 1. *Если для многошагового разностного метода (2) выполняется условие корней, то в векторном пространстве R^m можно определить векторную норму $||Z||$ такую, что согласованная с ней норма матрицы перехода S из (5) не превосходит единицу:*

$$||\, S\, || \;=\; \sup_{||Z|| \neq 0} \frac{||\, S\, Z\, ||}{||\, Z\, ||} \;\leq\; 1\,.$$

Доказательство. Пусть $\lambda_i \quad i = 1, 2, \dots, m$, – собственные значения матрицы S (корни характеристического полинома разностного метода (2)). Если жорданова форма $\tilde{S}$ матрицы S имеет диагональный вид, то через матрицу Q обозначим матрицу преобразования подобия такую, что $\tilde{S} = Q\, S\, Q^{-1}$. Заметим, что в этом случае $||\, \tilde{S}\, ||_\infty \leq 1$, так как на диагонали матрицы $\tilde{S}$ стоят собственные значения матрицы S, а по условию корней их модули не превосходят единицу.

Если матрица S имеет кратные собственные значения, то по условию корней, их модули строго меньше единицы и существует $\varepsilon > 0$ такое, что сумма модуля любого кратного собственного значения и ε будет строго меньше единицы. Тогда через матрицу Q обозначим матрицу преобразования подобия такую, что ненулевые внедиагональные элементы жордановой формы

$\tilde{S} = Q\,S\,Q^{-1}$ равны ε. Заметим, что и в этом случае $||\,\tilde{S}\,||_\infty \le 1$, так как сумма модулей элементов любой строки матрицы $\tilde{S}$ по построению и условию корней (равному по модулю единице собственному значению соответствует жорданов блок размерности 1) не превосходит единицу.

Определим в R^m норму $||\,Z\,|| = ||\,Q\,Z\,||_\infty$. Тогда

$$||\,S\,|| = \sup_{||Z||\neq 0} \frac{||\,S\,Z\,||}{||\,Z\,||} = \sup_{||QZ||_\infty \neq 0} \frac{||\,Q\,S\,Z\,||_\infty}{||\,Q\,Z\,||_\infty} =$$

$$= \sup_{||X||_\infty \neq 0} \frac{||\,\tilde{S}\,X\,||_\infty}{||\,X,||_\infty} = ||\,\tilde{S}\,||_\infty \le 1\,,$$

что и требовалось доказать. □

Теорема 1. *Пусть решение задачи Коши (1) определено и непрерывно на интервале* $[\,0,\,T\,]$, *а правая часть* $f(t,u)$ *определена и непрерывна при* $t \in [\,0,\,T\,]$, $u \in R$ *и, кроме того, непрерывна по Липшицу по второму аргументу во всей области своего определения.*

Пусть параметры многошагового разностного метода (2) не зависят от шага интегрирования τ.

Тогда существуют постоянные τ_0, C_1 *и* C_2 *такие, что для погрешности решения задачи (1) методом (2) при* $0 < \tau < \tau_0$ *имеет место оценка*

$$|\,z_{n+m-1}\,| \le C_1 \max_{0\le k\le m-1} |\,z_k\,| + C_2 \max_{m\le k\le n+m-1} |\,\psi_k\,|$$

для любого $t_{n+m-1} \in [\,0,\,T\,]$, *т.е. погрешность решения определяется точностью задания начальных данных и аппроксимации дифференциального уравнения.*

Доказательство. Пусть в R^m в соответствии с леммой 1 задана норма $||\,Z\,||$ и $||\,S\,|| \le 1$. Тогда из (5) следует, что

$$||\,Z^n\,|| \le ||\,Z^{n-1}\,|| + \frac{\tau}{|a_0|}\big(\,||\,\Phi^n\,|| + ||\,\Psi^n\,||\,\big)\,. \tag{7}$$

Так как в конечномерном пространстве R^m все нормы эквивалентны, то существуют положительные постоянные α и β такие, что

$$\alpha\,||\,Z\,||_\infty \le ||\,Z\,|| \le \beta\,||\,Z\,||_\infty \quad \forall\,Z \in R^m\,. \tag{8}$$

Заметим, что постоянные α и β не зависят от τ, поскольку определяющие норму $||\,Z\,||$ элементы матрицы S (параметры $a_0 \neq 0,\, a_1,\, \ldots\,,\, a_m$ разностного метода (2)) от него не зависят.

Тогда

$$\begin{aligned} ||\,\Phi^n\,|| &\le \beta\,||\,\Phi^n\,||_\infty = \beta\,|\,\varphi_n\,|\,,\\ ||\,\Psi^n\,|| &\le \beta\,||\,\Psi^n\,||_\infty = \beta\,|\,\psi_n\,|\,. \end{aligned} \tag{9}$$

Так как функция $f(t,u)$ непрерывна по Липшицу по второму аргументу с постоянной L, то

$$
\begin{aligned}
|\,\varphi_n\,| &= \Big|\,\sum_{k=0}^{m} b_k\,[\,f(t_{n-k}, y_{n-k}) - f(t_{n-k}, u(t_{n-k}))\,]\,\Big| \le \\
&\le |b_0|\cdot L\cdot|\,y_n - u(t_n)\,| + \sum_{k=1}^{m}|b_k|\cdot L\cdot|\,y_{n-k}-u(t_{n-k})\,| \le \\
&\le |b_0|\cdot L\cdot|\,z_n\,| + \sum_{k=1}^{m}|b_k|\cdot L\cdot|\,z_{n-k}\,| \le \\
&\le |b_0|\cdot L\cdot||\,Z^n\,||_\infty + \Big(\sum_{k=1}^{m}|b_k|\Big)\cdot L\cdot||\,Z^{n-1}\,||_\infty \le \\
&\le |b_0|\cdot L\cdot\frac{1}{\alpha}\,||\,Z^n\,|| + \Big(\sum_{k=1}^{m}|b_k|\Big)\cdot L\cdot\frac{1}{\alpha}\,||\,Z^{n-1}\,||\,.
\end{aligned}
$$

Из этой оценки, неравенств (7) и (9) следует, что

$$
||\,Z^n\,|| \le \tau\,\frac{\beta\,|b_0|\,L}{\alpha\,|a_0|}\,||\,Z^n\,|| + \Bigg[1 + \tau\,\frac{\beta\Big(\sum\limits_{k=1}^{m}|b_k|\Big)L}{\alpha\,|a_0|}\Bigg]\,||\,Z^{n-1}\,|| + \tau\,\frac{\beta}{|a_0|}\,|\,\psi_n\,|
$$

или

$$
\begin{gathered}
||\,Z^n\,|| \le \frac{1+\tau\,c_1}{1-\tau\,c_0}\,||\,Z^{n-1}\,|| + \tau\,\frac{c_2}{1-\tau\,c_0}\,|\,\psi_n\,|\,, \\
c_0 = \frac{\beta\,|b_0|\,L}{\alpha\,|a_0|}\,,\quad c_1 = \frac{\beta\Big(\sum\limits_{k=1}^{m}|b_k|\Big)L}{\alpha\,|a_0|}\,,\quad c_2 = \frac{\beta}{|a_0|}\,,
\end{gathered}
\tag{10}
$$

при $\tau \in (\,0,\, c_0^{-1}\,)$.

Заметим, что для явных многошаговых разностных методов $c_0 = 0$, следовательно, это неравенство будет выполняться при любом шаге интегрирования.

Из неравенства (10) следует, что

$$
\begin{aligned}
||\,Z^{n+m-1}\,|| &\le \frac{1+\tau\,c_1}{1-\tau\,c_0}\,||\,Z^{n+m-2}\,|| + \tau\,\frac{c_2}{1-\tau\,c_0}\,|\,\psi_{n+m-1}\,| \le \\
&\le \Big(\frac{1+\tau\,c_1}{1-\tau\,c_0}\Big)^2\,||\,Z^{n+m-3}\,|| + \\
&\qquad + \tau\,\frac{c_2}{1-\tau\,c_0}\Big[\frac{1+\tau\,c_1}{1-\tau\,c_0}\,|\,\psi_{n+m-1}\,| + |\,\psi_{n+m-2}\,|\Big] \le \\
&\le \Big(\frac{1+\tau\,c_1}{1-\tau\,c_0}\Big)^n\,||\,Z^{m-1}\,|| + \\
&\qquad + \tau\,\frac{c_2}{1-\tau\,c_0}\sum_{k=0}^{n-1}\Big(\frac{1+\tau\,c_1}{1-\tau\,c_0}\Big)^k\,|\,\psi_{n+m-1-k}\,| \le \\
&\le \Big(\frac{1+\tau\,c_1}{1-\tau\,c_0}\Big)^n\Big[\,||\,Z^{m-1}\,|| + \\
&\qquad + \frac{c_2}{c_0+c_1}\max_{m\le k\le n+m-1}|\,\psi_k\,|\Big]\,.
\end{aligned}
\tag{11}
$$

Так как для любого положительного x имеет место неравенство $(1+x)^{\frac{1}{x}} \le \mathrm{e}$, то

$$\left(\frac{1+\tau\, c_1}{1-\tau\, c_0}\right)^n = \left(1+\tau\,\frac{c_0+c_1}{1-\tau\, c_0}\right)^n \le \exp\left\{t_n\frac{c_0+c_1}{1-\tau\, c_0}\right\}$$

и, используя неравенства (9), из (11) легко получить следующую оценку погрешности решения в равномерной норме:

$$\begin{aligned} \|\, Z^{n+m-1}\,\|_\infty \le \alpha^{-1} \exp\left\{t_n\frac{c_0+c_1}{1-\tau\, c_0}\right\} \Big[\beta\,\|\, Z^{m-1}\,\|_\infty + \\ + \frac{c_2}{c_0+c_1} \max_{m\le k\le n+m-1} |\,\psi_k\,|\Big]. \end{aligned} \tag{12}$$

Из этого неравенства при любом $\tau \in (\,0,\, 0.5c_0^{-1}\,)$ следует оценка

$$\begin{aligned} |\, z_{n+m-1}\,| \le \alpha^{-1}\,\mathrm{e}^{2(c_0+c_1)T} \Big[\beta \max_{0\le k\le m-1} |\, z_k\,| + \\ + \frac{c_2}{c_0+c_1} \max_{m\le k\le n+m-1} |\,\psi_k\,|\Big] \end{aligned}$$

для любого $t_{n+m-1} \in [\,0,\, T\,]$. □

9.2. Выбор параметров многошагового разностного метода

Прежде всего отметим, что числовые параметры $a_0 \ne 0,\; a_1,\; \ldots,\; a_m$ и $b_0,\, b_1,\, \ldots,\, b_m$ разностного метода (2)

$$\frac{a_0 y_n + a_1 y_{n-1} + \ldots + a_m y_{n-m}}{\tau} = b_0 f_n + b_1 f_{n-1} + \ldots + b_m f_{n-m}\,,$$
$$n = m,\, m+1,\, \ldots,$$

определяются с точностью до постоянного множителя. Поэтому обычно на них накладывается условие нормировки

$$b_0 + b_1 + \ldots + b_m = 1\,. \tag{13}$$

Далее, если мы хотим найти параметры метода такие, при которых ошибка аппроксимации дифференциального уравнения

$$\psi_n = \frac{1}{\tau}\sum_{k=0}^{m} a_k\, u(t_{n-k}) - \sum_{k=0}^{m} b_k\, f(t_{n-k}, u(t_{n-k})) \tag{14}$$

была величиной $O(\tau^p)$, то разложим $u(t_{n-k}) = u(t_n - k\tau)$ и $f(t_{n-k}, u(t_{n-k})) = u'(t_n - k\tau)$ в ряды Тейлора в точке t_n:

$$u(t_{n-k}) = \sum_{i=0}^{p} u^{(i)}(t_n)\,\frac{(-k\,\tau)^i}{i\,!} + O(\,\tau^{p+1}\,)\,,$$
$$f(t_{n-k}) = \sum_{i=0}^{p-1} u^{(i+1)}(t_n)\,\frac{(-k\,\tau)^i}{i\,!} + O(\,\tau^{p}\,)$$

и, подставляя эти разложения в (14), получим

$$\psi_n = \Big[\, \sum_{k=0}^{m} a_k \Big] \frac{u(t_n)}{\tau} + \\ + \sum_{i=1}^{p} \Big[\sum_{k=0}^{m} \big(a_k\, k^i \,+\, b_k\, i\, k^{i-1} \big) \Big] u^{(i)}(t_n)\, \frac{(-\tau)^i}{(i-1)\,!} \,+\, O(\,\tau^p\,)\,. \tag{15}$$

Следовательно, если параметры многошагового разностного метода выбрать так, чтобы суммы в квадратных скобках представления (15) обратились в нуль, т.е.

$$\begin{aligned} &\sum_{k=0}^{m} a_k \,=\, 0\,, \\ &\sum_{k=0}^{m} \big(a_k\, k^i \,+\, b_k\, i\, k^{i-1} \big) \,=\, 0\,, \quad i = 1, 2, \ldots, p\,, \\ &\sum_{k=0}^{m} b_k \,=\, 1\,, \end{aligned} \tag{16}$$

(здесь мы добавили условие нормировки (13)), то ошибка аппроксимации будет величиной $O(\tau^p)$, а сами параметры не будут зависеть от шага интегрирования τ , так как система (16) от него не зависит. Последний факт позволяет применять теорему 1 для оценки погрешности решения, если выполняется условие корней. □

Теперь нам нужно выяснить, когда система (16) имеет решение и $a_0 \neq 0$. Перепишем систему (16) (без последнего уравнения) в следующем виде:

$$\sum_{k=0}^{m} \big[\, a_k\, x^i \,+\, b_k\, (x^i)'\, \big]_{x=k} \,=\, 0\,, \quad i = 0, 1, \ldots, p\,. \tag{17}$$

Умножим i-тое уравнение на произвольную постоянную q_i и все уравнения просуммируем:

$$\sum_{k=0}^{m} \big[\, a_k\, (q_0 + q_1 x + \ldots + q_p x^p) \,+\, b_k\, (q_0 + q_1 x + \ldots + q_p x^p)'\, \big]_{x=k} \,=\, 0\,,$$

т.е. параметры m-шагового разностного метода (2) с ошибкой аппроксимации $O(\tau^p)$ удовлетворяют тождеству

$$\sum_{k=0}^{m} \big[\, a_k\, Q_p(k) \,+\, b_k\, Q_p'(k)\, \big] \,=\, 0 \tag{18}$$

для любого полинома $Q_p(x)$ степени p.

Вспомним, что задача построения полинома $Q_p(x)$, удовлетворяющего условиям:

$$Q_p(k) = \alpha_k\,, \quad Q_p'(k) = \beta_k\,, \quad k = 0, 1, \ldots, m\,,$$

(задача интерполирования с кратными узлами) имеет решение при любых $\alpha_0,\ \alpha_1,\ \ldots,\ \alpha_m$ и $\beta_0,\ \beta_1,\ \ldots,\ \beta_m$, если $p \geq 2m+1$.

Следовательно, при $p \geq 2m+1$ система (17) имеет только нулевое решение, так как, задав $\alpha_k = a_k$ и $\beta_k = b_k$, получим

$$\sum_{k=0}^{m} \left(a_k^2 + b_k^2 \right) = 0\,,$$

что возможно только в случае $a_k = 0$ и $b_k = 0$, $k = 0, 1, \ldots, m$. Тем самым доказана

Лемма 2. *Система (16) не имеет решения при $p \geq 2m+1$, т.е. не существует m-шаговых разностных методов с погрешностью аппроксимации $O(\tau^{2m+1})$.* □

Заметим, что строки матрицы системы (17) линейно независимы при $p \leq 2m+1$, так как при $p = 2m+1$ определитель матрицы отличен от нуля (однородная система имеет только нулевое решение). Следовательно, при $p \leq 2m$ ранг матрицы системы (17) равен p (все ее строки линейно независимы), ее общее решение принадлежит $(2m+1-p)$-мерному ядру матрицы, из которого выбирается решение удовлетворяющее условию нормировки (13) и, если это возможно, условию корней, обеспечивающему устойчивость вычислений.

9.3. Неявные разностные методы максимального порядка аппроксимации

Получим формулы для решения системы (17) при $p = 2m$. Для этого определим полином степени $m+1$

$$W(x) = x\,(x-1)\,\ldots\,(x-m)$$

и полиномы степени $2m$

$$B_i(x) = \frac{W^2(x)}{x(x-i)}\,, \quad i = 1\,, \ldots, m\,.$$

Тогда, из (18) при $Q_{2m}(x) = B_i(x)$ получим, что

$$B_i'(0)\,b_0 + B_i'(i)\,b_i = 0\,, \quad i = 1\,, \ldots, m\,. \tag{19}$$

Так как нетрудно показать, что

$$B_i'(0) = -\frac{\left[W'(0)\right]^2}{i} = -\frac{(m\,!)^2}{i}\,,$$
$$B_i'(i) = \frac{\left[W'(i)\right]^2}{i} = \frac{[i\,!\,(m-i)\,!]^2}{i}\,, \quad i = 1\,, \ldots, m\,,$$

то из (19) получаем, что

$$b_i = \left(\frac{m!}{i!\,(m-i)!} \right)^2 b_0 \,, \quad i = 1\,, \ldots, m\,. \tag{20}$$

Для вычисления параметров a_i определим полиномы степени $2m$

$$A_i(x) = \frac{W^2(x)}{(x-i)^2}\,, \quad i = 0,\,1,\,\ldots,\,m\,.$$

Тогда, из (18) при $Q_{2m}(x) = A_i(x)$ получим, что

$$A_i(i)\,a_i + A_i'(i)\,b_i = 0\,, \quad i = 0,\,1,\,\ldots,\,m\,. \tag{21}$$

Так как нетрудно показать, что

$$A_i(i) = \big[\,W'(i)\,\big]^2, \quad A_i'(i) = 2\,\big[\,W'(i)\,\big]^2 \sum_{k=0,\,k\neq i}^{m} \frac{1}{i-k}\,,$$
$$i = 0,\,1,\,\ldots,\,m\,,$$

то из (21) получаем, что

$$a_i = 2\Big(\sum_{k=0,\,k\neq i}^{m} \frac{1}{k-i} \Big)\,b_i\,, \quad i = 0,\,1,\,\ldots,\,m\,. \tag{22}$$

Из формул (20) и (22) следует, что параметры m-шагового разностного метода определяются сс точностью до множителя b_0, который можно выбрать из условия нормировки (13), и $a_0 \neq 0$. Следовательно, доказана

Теорема 2. *Существует неявный m-шаговый разностный метод с погрешностью аппроксимации $O(\tau^{2m})$, его параметры определяются по формулам (20) и (21). Явных m-шаговых разностных методов с погрешностью аппроксимации $O(\tau^{2m})$ не существует.* □

При $m = 1$ получим одношаговый метод порядка $O(\tau^2)$:

$$\frac{y_n - y_{n-1}}{\tau} = \frac{f_n + f_{n-1}}{2}\,. \tag{23}$$

При $m = 2$ получим двухшаговый метод порядка $O(\tau^4)$:

$$\frac{y_n - y_{n-2}}{2\tau} = \frac{f_n + 4f_{n-1} + f_{n-2}}{6}\,. \tag{24}$$

Нетрудно убедиться, что эти методы удовлетворяют условию корней, но при $m = 3$ трехшаговый метод порядка $O(\tau^6)$:

$$\frac{11y_n + 27y_{n-1} - 27y_{n-2} - 11y_{n-3}}{60\tau} = \frac{f_n + 9f_{n-1} + 9f_{n-2} + f_{n-3}}{20}\,, \tag{25}$$

ему не удовлетворяет, так как один из корней характеристического уравнения равен $\frac{-19-\sqrt{241}}{11} < -1$.

9.4. Явные разностные методы максимального порядка аппроксимации

Получим формулы для решения системы (17) при $p = 2m - 1$ и $b_0 = 0$. Для этого определим полином степени m

$$w(x) \,=\, (x-1)\,(x-2)\,\ldots\,(x-m)$$

и полиномы степени $2m-1$

$$B_i(x) \,=\, \frac{w^2(x)}{x-i}\,, \quad i \,=\, 1\,,\,\ldots\,,\,m\,.$$

Тогда, из (18) при $Q_{2m-1}(x) = B_i(x)$ получим, что

$$B_i(0)\,a_0 \,+\, B_i'(i)\,b_i \,=\, 0\,, \quad i \,=\, 1\,,\,\ldots\,,\,m\,. \tag{26}$$

Так как нетрудно показать, что

$$\begin{aligned} B_i(0) \,&=\, -\,\frac{\big[\,w(0)\,\big]^2}{i} \,=\, -\,\frac{(\,m\,!\,)^2}{i}\,, \\ B_i'(i) \,&=\, \big[\,w'(i)\,\big]^2 \,=\, [\,(i-1)\,!\,(m-i)\,!\,]^2\,, \quad i \,=\, 1\,,\,\ldots\,,\,m\,, \end{aligned}$$

то из (26) получаем, что

$$b_i \,=\, \frac{1}{i}\left(\frac{m\,!}{(i-1)\,!\,(m-i)\,!}\right)^2 a_0\,, \quad i \,=\, 1\,,\,\ldots\,,\,m\,. \tag{27}$$

Для вычисления параметров a_i определим полиномы степени $2m-1$

$$A_i(x) \,=\, \frac{x\,w^2(x)}{(x-i)^2}\,, \quad i \,=\, 1,\,\ldots\,,\,m\,.$$

Тогда, из (18) при $Q_{2m-1}(x) = A_i(x)$ получим, что

$$A_i(i)\,a_i \,+\, A_i'(i)\,b_i \,=\, 0\,, \quad i \,=\, 1,\,\ldots\,,\,m\,. \tag{28}$$

Так как нетрудно показать, что

$$A_i(i) \,=\, i\,\big[\,w'(i)\,\big]^2, \quad A_i'(i) \,=\, \big[\,w'(i)\,\big]^2\left[\,1 \,+\, 2\,i\sum_{k=1,\,k\neq i}^{m}\frac{1}{i-k}\,\right],$$
$$i \,=\, 1,\,\ldots\,,\,m\,,$$

то из (28) получаем, что

$$a_i \,=\, -\Big(\,\frac{1}{i} \,+\, 2\sum_{k=1,\,k\neq i}^{m}\frac{1}{i-k}\,\Big)\,b_i\,, \quad i \,=\, 1,\,\ldots\,,\,m\,. \tag{29}$$

Из формул (27) и (29) следует, что параметры явного m-шагового разностного метода определяются с точностью до множителя a_0, который можно выбрать из условия нормировки (13), и $a_0 \neq 0$. Следовательно, доказана

Теорема 3. *Существует явный m-шаговый разностный метод с погрешностью аппроксимации $O(\tau^{2m-1})$, его параметры определяются по формулам (27) и (29).* □

При $m = 1$ получим одношаговый метод порядка $O(\tau)$:

$$\frac{y_n - y_{n-1}}{\tau} = f_{n-1}, \tag{30}$$

удовлетворяющий условию корней, но уже при $m = 2$ двухшаговый метод порядка $O(\tau^3)$:

$$\frac{y_n + 4y_{n-1} - 5y_{n-2}}{6\tau} = \frac{2f_{n-1} + f_{n-2}}{3}, \tag{31}$$

ему не удовлетворяет, так как один из корней характеристического уравнения равен $-5 < -1$.

9.5. Методы Адамса

Очевидно, что условию корней удовлетворяют m-шаговые разностные методы

$$\frac{y_n - y_{n-1}}{\tau} = b_0 f_n + b_1 f_{n-1} + \ldots + b_m f_{n-m}, \tag{32}$$

с ошибкой аппроксимации $O(\tau^{m+1})$ (неявный метод Адамса), если параметры $\{b_i\}$ определить по формулам (в основе лежит приближение правой части дифференциального уравнения по ее значениям в узлах t_{n-m}, t_{n-m+1}, $\ldots$, t_{n-1}, t_n интерполяционным полиномом)

$$b_i = \int\limits_0^1 \frac{W(x)}{(x-i)\,W'(i)}\,dx, \quad i = 0, 1, \ldots, m,$$

с ошибкой аппроксимации $O(\tau^m)$ (явный метод Адамса), если параметры $\{b_i\}$ определить по формулам (в основе лежит приближение правой части дифференциального уравнения по ее значениям в узлах t_{n-m}, t_{n-m+1}, $\ldots$, t_{n-1} интерполяционным полиномом)

$$b_0 = 0, \quad b_i = \int\limits_0^1 \frac{w(x)}{(x-i)\,w'(i)}\,dx, \quad i = 1, \ldots, m.$$

10. Численное решение интегральных уравнений

Уравнение вида

$$F(x,\, u(x),\, y) \;=\; 0\,,$$

$$y = \int\limits_a^x K(\,x, s, u(s)\,)\, ds \qquad y = \int\limits_a^b K(\,x, s, u(s)\,)\, ds\,,$$

относительно неизвестной на интервале $[\,a,\, b\,]$ функции $u(x)$ называется интегральным уравнением. Простейшими представителями этого класса являются линейное интегральное уравнение Вольтерра II рода:

$$u(x) \,+\, \int\limits_a^x K(x,s)\, u(s)\, ds \;=\; f(x)\,, \tag{1}$$

и линейное интегральное уравнение Фредгольма II рода:

$$u(x) \,+\, \int\limits_a^b K(x,s)\, u(s)\, ds \;=\; f(x)\,, \tag{2}$$

где функции $K(x,s)$ (ядро уравнения) и $f(x)$ (правая часть уравнения) заданы.

Уравнения Вольтерра и Фредгольма, вид которых отличается лишь верхним пределом интеграла, принципиально различны: решение первого в точке x зависит от своих значений только в предыдущих точках, что аналогично случаю систем линейных алгебраических уравнений с нижней треугольной матрицей, а решение уравнения Фредгольма в точке x зависит от своих значений во всех точках интервала $[\,a,\, b\,]$ что аналогично случаю систем линейных алгебраических уравнений с плотной матрицей, которые решать значительно сложнее.

Мы будем предполагать, что ядра и правые части интегральных уравнений (1) и (2) по крайней мере непрерывны, тогда уравнения имеют непрерывные решения, что устанавливается с помощью принципа сжимающих отображений.

Теорема 1. *Если функция $K(x,y)$ непрерывна в квадрате $a \le x, y \le b$, а функция $f(x)$ непрерывна на интервале $[\,a,\, b\,]$, то решение интегрального уравнения Вольтерра*

$$u(x) \,+\, \int\limits_a^x K(x,s)\, u(s)\, ds \;=\; f(x)$$

в классе непрерывных функций существует и единственно.

Доказательство. Пусть $K_0 = \max\limits_{a \le x,s \le b} |\, K(x,s)\,|$, а положительное целое число m выбрано так, чтобы $q = dK_0 < 1$, где $d = \frac{(b-a)}{m}$.

Определим отображение φ из $C\,[\,a,\, a+d\,]$ в $C\,[\,a,\, a+d\,]$:

$$\varphi(v) \;=\; f(x) \;-\; \int\limits_a^x K(x,s)\, v(s)\, ds\, .$$

Так как

$$|| \, \varphi(v) - \varphi(w) \, ||_{C[a,a+d]} = \max_{a \le x \le a+d} \left| \, \int\limits_a^x K(x,s) \left[\, w(s) - v(s) \,\right] ds \, \right| \le$$
$$\le q \, || \, v - w \, ||_{C[a,a+d]} \, ,$$

а $q < 1$, то отображение φ является сжимающим и, следовательно, уравнение $u = \varphi(u)$ (уравнение Вольтерра) имеет в $C\,[\,a,\, a+d\,]$ единственное решение.

Для окончательного доказательства теоремы воспользуемся методом математической индукции. Предположим, что решение уравнения Вольтерра на интервале $[\,a,\, a+kd\,]$, $k < m$, существует и непрерывно. Определив отображение

$$\varphi_k(v) \;=\; f_k(x) \;-\; \int\limits_{a+kd}^x K(x,s)\, v(s)\, ds$$

с непрерывной функцией $f_k(x) = f(x) - \int\limits_a^{a+kd} K(x,s)\, v(s)\, ds$ и повторив аналогичные случаю $k = 1$ рассуждения, установим, что $\varphi_k(v)$ является сжимающим отбражением из $C\,[\,a+kd,\, a+(k+1)d\,]$ в $C\,[\,a+kd,\, a+(k+1)d\,]$ и, следовательно, уравнение $u_k = \varphi_k(u_k)$ (уравнение Вольтерра) имеет в $C\,[\,a+kd,\, a+(k+1)d\,]$ единственное решение, которое (это практически очевидно) является непрерывным продолжением решения $u(x)$ уравнения Вольтерра на интервале $[\,a,\, a+kd\,]$. Так как для $k = 1$ предположение матиндукции доказано, то оно верно и для $k = m$. □

Теорема 2. *Если функция $f(x)$ непрерывна на интервале $[\,a,\, b\,]$, функция $K(x,y)$ непрерывна в квадрате $a \le x,y \le b$ и $\max\limits_{a \le x,s \le b} |\, K(x,s)\,| < 1$, то решение интегрального уравнения Фредгольма*

$$u(x) \;+\; \int\limits_a^b K(x,s)\, u(s)\, ds \;=\; f(x)\, ,$$

в классе непрерывных функций существует и единственно.

Доказательство этой теоремы совпадает с первой частью доказательства теоремы 1 . □

10.1. Метод квадратур решения уравнения Вольтерра II рода

Решение линейного интегрального уравнения Вольтерра II рода:

$$u(x) + \int\limits_a^x K(x,s)\,u(s)\,ds = f(x) \tag{1}$$

будем искать в узлах равномерной сетки

$$x_0 = a\,, \quad x_1 = a+h\,, \quad \ldots\,, \quad x_n = a+nh\,, \quad h = \frac{b-a}{n}\,. \tag{3}$$

Для определения неизвестных значений $\{u(x_i)\}_{i=0}^n$ перепишем уравнение (1) при $x = x_i$:

$$u(x_0) = f(x_0)$$

$$u(x_i) + \int\limits_a^{x_i} K(x_i,s)\,u(s)\,ds = f(x_i)\,, \quad i = 1, \ldots, n\,.$$

Интеграл по отрезку $[\,a,\, x_i\,]$ заменим квадратурной формулой на узлах x_0, x_1, …, x_i с весами hA_{ij} :

$$\begin{gathered} u(x_0) = f(x_0) \\ u(x_i) + h\sum_{j=0}^{i} A_{ij}\,K(x_i,x_j)\,u(x_j) = f(x_i) - \psi_i\,, \quad i = 1, \ldots, n\,, \end{gathered} \tag{4}$$

где ψ_i – погрешность квадратурной формулы, зависящая от известной функции $K(x_i,x)$ и неизвестного решения $u(x)$.

Считая погрешности ψ_i являются малыми величинами и отбрасывая их получаем систему линейных алгебраических уравнений

$$\begin{gathered} u_0 = f(x_0) \\ u_i + h\sum_{j=0}^{i} A_{ij}\,K(x_i,x_j)\,u_j = f(x_i)\,, \quad i = 1, \ldots, n\,, \end{gathered} \tag{5}$$

относительно неизвестных $\{u_i\}_{i=0}^n$, приближающих искомые значения решения $\{u(x_i)\}_{i=0}^n$ в узлах сетки.

Определение 1. Построение и решение системы (5) называется методом квадратур (или вычислительным правилом) приближенного решения интегрального уравнения Вольтерра (1). □

Очевидно, что матрица системы (5) является нижней треугольной матрицей и условием ее невырожденности является отличие от нуля диагональных элементов $1 + h\, K(x_i, x_i)\, A_{ii}$.

Обычно веса квадратурных формул (зависящие от количества сеточных узлов) выбирают так, чтобы

$$\max_{1\le i\le n}\Big[\max_{0\le j\le i}|\,A_{ij}(n)\,|\Big] \;\le\; A \;<\; \infty \quad \forall\, n\,. \tag{6}$$

Это условие гарантирует применимость метода квадратур при достаточно малом шаге сетки.

Теорема 3. *Линейная алгебраическая система (5) метода квадратур приближенного решения интегрального уравнения Вольтерра (1) с условием (6) имеет единственное решение при любом* $n \;\ge\; A\,K_0 \,+\, 1$, *где* $K_0 = \max\limits_{a\le x,s\le b}|\,K(x,s)\,|$.

Справедливость этого утверждения следует из того, что при выполнении условий теоремы диагональные элементы матрицы системы (5) строго положительны. □

Пример. Если в качестве квадратур взять составные формулы трапеций, то получим метод квадратур

$$\begin{aligned} u_0 \;&=\; f(x_0)\,,\\ u_i \,+\, h\,\big[\,0.5\,K(x_i,x_0)\,u_0 \,+\, K(x_i,x_1)\,u_1 \,+\, \ldots \,+\\ +\,K(x_i,x_{n-1})\,u_{n-1} \,+\, 0.5\,K(x_i,x_n)\,u_n\,\big] \;&=\; f(x_i)\,,\\ i \,=\, 1,\,\ldots\,,\,n\,,& \end{aligned}$$

для которого условие (6) выполняется. □

Условие (6) позволяет оценить погрешность решения

$$\varepsilon_i \;=\; u_i \,-\, u(x_i)\,, \quad i \,=\, 0,1,\,\ldots\,,\,n\,, \tag{7}$$

через погрешности квадратурных формул при достаточно малом шаге сетки.

Теорема 4. *Если для метода квадратур (5) приближенного решения интегрального уравнения Вольтерра (1) выполняется условие (6) и* $q = h\,A\,K_0 \le 0.5$, *где* $K_0 = \max\limits_{a\le x,s\le b}|\,K(x,s)\,|$, *то для погрешности решения (7) справедлива оценка*

$$|\,\varepsilon_i\,| \;\le\; \mathrm{e}^{2\,(b-a)\,A\,K_0}\,\psi\,, \quad i \,=\, 1,\,\ldots\,,\,n\,, \tag{8}$$

где $\psi \;=\; \max\limits_{1\le i\le n}|\,\psi_i\,|$ *– максимальная погрешность квадратурных формул метода.*

Доказательство. Вычтем из (5) соотношения (4):

$$\varepsilon_0 = 0\,, \quad \varepsilon_i + h\sum_{j=0}^{i} A_{ij}\,K(x_i,x_j)\,\varepsilon_j = \psi_i\,, \quad i = 1, \ldots, n\,.$$

Отсюда, так как $h\,|\,A_{ij}\,K(x_i,x_j)\,| \le q = h\,A\,K_0 < 1$ для любых $i = 1, \ldots, n$, $j = 0, 1, \ldots, i$, следует, что

$$|\,\varepsilon_i\,| \le \frac{\psi}{1-q} + \frac{q}{1-q}\sum_{j=0}^{i-1}|\,\varepsilon_j\,| \quad i = 1, \ldots, n\,.$$

Определим последовательность неотрицательных чисел

$$E_0 = |\,\varepsilon_0\,| = 0\,, \quad E_i = \frac{\psi}{1-q} + \frac{q}{1-q}\sum_{j=0}^{i-1}E_j\,, \quad i = 1, \ldots, n\,.$$

Очевидно, что $|\,\varepsilon_i\,| \le E_i$ для любого $i = 1, \ldots, n$. Далее, так как

$$E_i = \Big[\frac{\psi}{1-q} + \frac{q}{1-q}\sum_{j=0}^{i-2}E_j\Big] + \frac{q}{1-q}\,E_{i-1} = \frac{1}{1-q}\,E_{i-1}\,,$$

то $E_i = (1-q)^{1-i}\,E_1 = (1-q)^{-i}\,\psi$ и, следовательно,

$$|\,\varepsilon_i\,| \le \Big(\frac{1}{1-q}\Big)^i\psi \le \Big(1+\frac{q}{1-q}\Big)^n\psi \le \mathrm{e}^{\frac{q\,n}{1-q}}\,\psi\,,$$

откуда и следует оценка (8) при $q \le 0.5$. □

10.2. Метод квадратур решения уравнения Фредгольма II рода

Запишем линейное интегральное уравнение Фредгольма II рода в виде

$$u(x) - \lambda\int_a^b K(x,s)\,u(s)\,ds = f(x)\,, \tag{9}$$

где параметр λ выделен специально: при некоторых его значениях, например, $|\,\lambda\,| < \big[\,(b-a)\max\limits_{a\le x,s\le b}|\,K(x,s)\,|\,\big]^{-1}$ решение существует и единственно.

На отрезке $[\,a,\,b\,]$ зададим сетку из n узлов

$$a \le x_1 < x_2 < \ldots < x_n \le b \tag{10}$$

и будем предполагать, что при любом $x \in [\,a,\,b\,]$ для вычисления интеграла из (9) известна квадратурная формула

$$\int_a^b K(x,s)\,u(s)\,ds \;=\; \sum_{i=1}^{n} A_i(x)\,u(x_i) \;+\; \psi(x) \tag{11}$$

с весами $A_i(x)$ и погрешностью $\psi(x)$.

Для определения неизвестных значений $\{u(x_k)\}_{k=1}^{n}$ перепишем уравнение (9) при $x = x_k$ с учетом тождества (11):

$$\begin{gathered} u(x_k) \;-\lambda \sum_{i=1}^{n} A_i(x_k)\,u(x_i) \;=\; f(x_k) \;+\; \lambda\,\psi(x_k)\,, \\ k \;=\; 1,\,\ldots\,,\,n\,. \end{gathered} \tag{12}$$

Считая погрешности аппроксимации уравнения $\lambda\,\psi(x_k)$ малыми величинами и отбрасывая их, получаем систему линейных алгебраических уравнений

$$\begin{gathered} u_k \;-\lambda \sum_{i=1}^{n} A_i(x_k)\,u_i \;=\; f(x_k)\,, \\ k \;=\; 1,\,\ldots\,,\,n\,, \end{gathered} \tag{13}$$

относительно неизвестных $\{u_k\}_{k=0}^{n}$, приближающих искомые значения решения $\{u(x_k)\}_{k=0}^{n}$ в узлах сетки.

Определение 2. Построение и решение системы (13) называется методом квадратур приближенного решения интегрального уравнения Фредгольма (9). □

Матрицу системы (13) представима в виде $E - \lambda\,A$, где E – единичная матрица порядка n и C – матрица весов квадратурных формул:

$$A = \begin{bmatrix} A_1(x_1) & A_2(x_1) & \ldots & A_n(x_1) \\ A_1(x_2) & A_2(x_2) & \ldots & A_n(x_2) \\ \cdots & \cdots & \cdots & \cdots \\ A_1(x_n) & A_2(x_n) & \ldots & A_n(x_n) \end{bmatrix}.$$

Обычно веса $A_i(x) \;=\; A_i^{(n)}(x)$ квадратурных формул (11) на семействе узлов

$$a \;\le\; x_1^{(n)} \;<\; x_2^{(n)} \;<\; \ldots \;<\; x_n^{(n)} \;\le\; b$$

выбирают так, чтобы

$$||\,A\,||_\infty \;=\; ||\,A^{(n)}\,||_\infty \;\le\; C \;<\; \infty \quad \forall\, n \ge 1\,. \tag{14}$$

Это условие позволяет указать интервал значений параметра λ, при которых система (13) всегда имеет решение.

Теорема 5. *Если для метода квадратур (13) выполняется условие (14), то система линейных алгебраических уравнений (13) имеет единственное решение при* $\lambda \in (\,-C^{-1},\,C^{-1}\,)$.

Доказательство. При $\lambda = 0$ система (13) очевидным образом имеет единственное решение.

Предположим, что $\lambda \in (\,-C^{-1},\, C^{-1}\,)$ и $\lambda \neq 0$, но система (13) либо не имеет решения либо имеет несколько решений, т.е. определитель матрицы $E - \lambda\, A$ равен нулю.

Следовательно, λ^{-1} – собственное значение матрицы A и ему соответствует ненулевой собственный вектор $v = \lambda\, A\, v$, тогда

$$||\, v\,||_\infty \;=\; |\,\lambda\,| \cdot ||\, A\, v\,||_\infty \;\le\; |\,\lambda\,|\, C\, ||\, v\,||_\infty\,,$$

т.е. $|\,\lambda\,| \ge C^{-1}$, что противоречит условию на λ. □

Пример. Если в качестве квадратур взять составные формулы прямоугольников с центральными точками (узлами)

$$x_1^{(n)} = a + 0.5h,\;\; x_2^{(n)} = a + 1.5h,\;\; \ldots,\;\; x_n^{(n)} = a + (n - 0.5)h\,, \quad h = \frac{b-a}{n}\,,$$

то получим метод квадратур

$$\begin{gathered} u_k \;-\; \lambda\,\bigl[\,h\,K(x_k, x_1)\,u_1 \;+\; h\,K(x_k, x_2)\,u_2 \;+\; \ldots \;+ \\ +\,h\,K(x_k, x_{n-1})\,u_{n-1} \;+\; h\;K(x_k, x_n)\,u_n\,\bigr] \;=\; f(x_k)\,, \\ k \;=\; 1, \,\ldots\,,\, n\,, \end{gathered}$$

с весами $A_i(x) = h\,K(x, x_i)$, для которого условие (14) выполняется с постоянной $C = (b-a) \max\limits_{a \le x,s \le b} |\,K(x,s)\,|$. □

Условие (14) позволяет оценить погрешность решения через погрешности квадратурных формул при достаточно малом параметре λ.

Теорема 6. *Если для метода квадратур (13) приближенного решения интегрального уравнения Фредгольма (9) выполняется условие (14) и* $\lambda \in (\,-C^{-1},\, C^{-1}\,)$, *то для погрешности решения справедлива оценка*

$$\begin{aligned} |\,u(x_k) \;-\; u_k)\,| \le |\,\lambda\,| \cdot ||\,(E - \lambda A)^{-1}\,||_\infty \max_{1 \le k \le n} |\,\psi(x_k)\,| \;\le \\ \le \frac{|\,\lambda\,|}{1 - |\,\lambda\,|\,C} \max_{1 \le k \le n} |\,\psi(x_k)\,|\,, \quad k \;=\; 1, \,\ldots\,,\, n\,. \end{aligned} \tag{15}$$

Доказательство. Прежде всего заметим, что из теоремы 5 следует существование матрицы $(E - \lambda A)^{-1}$.

Тогда первое неравенство из (15) является следствием того, что погрешность решения $\varepsilon_k = u(x_k) - u_k$ удовлетворяет системе линейных алгебраических уравнений

$$\varepsilon_k \;-\; \lambda \sum_{i=1}^{n} A_i(x_k)\,\varepsilon_i \;=\; \lambda\,\psi(x_k)\,, \quad k \;=\; 1, \,\ldots\,,\, n\,,$$

получаемой вычитанием уравнений (13) из равенств (12), с матрицей $E - \lambda A$.

Второе неравенство из (15) является следствием оценки

$$\|\,(E-\lambda A)^{-1}\,\|_\infty = \sup_{v\neq 0}\frac{\|\,(E-\lambda A)^{-1}v\,\|_\infty}{\|\,v\,\|_\infty} = \sup_{w\neq 0}\frac{\|\,w\,\|_\infty}{\|\,(E-\lambda A)w\,\|_\infty} \le$$
$$\le \sup_{w\neq 0}\frac{\|\,w\,\|_\infty}{(1-|\,\lambda\,|\cdot\|\,A\,\|)\,\|\,w\,\|_\infty} \le \frac{1}{1-|\,\lambda\,|\,C},$$

так как $|\,\lambda\,|\,C<1$ по условию теоремы. □

Замечание. Из доказательства теоремы 6 следует, что первое неравенство оценки (15) выполняется при любом λ, лишь бы существовала матрица $(E-\lambda A)^{-1}$. □

10.3. Апостериорная оценка погрешности квадратур

В оценку (15) погрешности решения уравнения Фредгольма (9) методом квадратур (13) входит неизвестная величина

$$\max_{1\le k\le n}|\,\psi(x_k)\,|, \qquad \psi(x) = \int_a^b K(x,s)\,u(s)\,ds - \sum_{i=1}^n A_i(x)\,u(x_i)$$

– погрешность квадратур (11) зависит от неизвестного решения $u(x)$ интегрального уравнения. Перепишем выражение для погрешности квадратур, заменив функцию $u(x)$ суммой правой части и интегрального члена уравнения (9):

$$\begin{aligned}\psi(x) = &\int_a^b K(x,s)\Big[\,f(s) + \lambda\int_a^b K(s,y)\,u(y)\,dy\Big]\,ds \\ &- \sum_{i=1}^n A_i(x)\Big[\,f(x_i) + \lambda\int_a^b K(x_i,y)\,u(y)\,dy\Big].\end{aligned} \tag{16}$$

Из правой части этого равенства легко выделить слагаемое

$$\psi_f(x) = \int_a^b K(x,s)\,f(s)\,ds - \sum_{i=1}^n A_i(x)\,f(x_i), \tag{17}$$

содержащее только известные величины и являющееся погрешностью квадратуры (11), примененной к функции $f(s)$.

Остаток правой части равенства (16) можно переписать в виде

$$\lambda\int_a^b\Big[\int_a^b K(x,s)\,K(s,y)\,ds - \sum_{i=1}^n A_i(x)\,K(x_i,y)\Big]\,u(y)\,dy,$$

где выражение в квадратных скобках

$$\psi_K(x,y) \;=\; \int\limits_a^b K(x,s)\,K(s,y)\,ds \;-\; \sum_{i=1}^{n} A_i(x)\,K(x_i,y) \tag{18}$$

содержит только известные величины и является погрешностью квадратуры (11), примененной к функции $K(s,y)$ при любом $y \in [\,a,\,b\,]$.

Тогда, так как $\psi(x) = \psi_f(x) + \lambda \int\limits_a^b \psi_K(x,y) u(y)\,dy$, то

$$|\,\psi(x)\,| \;\le\; |\,\psi_f(x)\,| \;+\; |\,\lambda\,| \cdot ||\,u(y)\,||_{C[a,b]} \int\limits_a^b |\,\psi_K(x,y)\,|\,dy\,. \tag{19}$$

Эту оценку погрешности квадратур (11) можно вычислить в узлах сетки, заменив норму решения максимальным по модулю значением приближенного решения.

Замечание. Так как из интегрального уравнения (9) следует, что

$$||\,u(x)\,||_{C[a,b]} \;\le\; ||\,f(x)\,||_{C[a,b]} \;+\; |\,\lambda\,|\,(b-a)\,||\,K(x,s)\,||_{C([a,b]\times[a,b])}\,,$$

то при $|\,\lambda\,| < \big[\,(b-a) \max\limits_{a\le x,s\le b} |\,K(x,s)\,|\,\big]^{-1}$ имеет место оценка

$$||\,u(x)\,||_{C[a,b]} \;\le\; \frac{||\,f(x)\,||_{C[a,b]}}{1 - |\,\lambda\,|\,(b-a)\,||\,K(x,s)\,||_{C([a,b]\times[a,b])}}\,,$$

с учетом которой оценка погрешности квадратур (11) становится гарантированной.

Список литературы

[1] *Бабенко К.И.* Основы численного анализа.- М.: Наука, 1986.

[2] *Бахвалов Н.С.* Численные методы. - М.: Наука, 1975.

[3] *Бахвалов Н.С., Жидков Н.П., Кобельков Г.М.* Численные методы. - М.: Наука, 1987.

[4] *Березин И.С., Жидков Н.П.* Методы вычислений. - Ч.1.- М.: Наука, 1966. То же.- Ч.2.- Физматгиз, 1962

[5] *Калиткин Н.Н.* Численные методы. - М.: Наука, 1978.

[6] *Крылов В.И., Бобков В.В., Монастырный П.И.* Вычислительные методы.- Т.I.- М.: Наука, 1976. То же.- Т.II.- М.: Наука, 1977

[7] *Самарский А.А., Гулин А.В.* Численные методы. - М.: Наука, 1989.