Математика и математическое моделирование. 2019. № 5. С. 32–48.
DOI: 10.24108/mathm.0519.0000205

Математика и Математическое моделирование
Сетевое научное издание
http://mathmelpub.ru ISSN 2412-5911

УДК 004.3+519.6

Анализ погрешности некоторых явных конечно-разностных методов решения задачи Коши на примере модельного уравнения Далквиста

Ахрем А. А.[1], Носов А. П.[1,*], Рахманкулов В. З.[1], Южанин К. В.[1]

[1]Федеральный исследовательский центр "Информатика и управление" Российской Академии Наук, Москва, Россия
[*]nosov@isa.ru

В работе проведено исследование погрешности явных конечно-разностных методов Рунге — Кутты первого и второго порядков точности решения задачи Коши модельного уравнения Далквиста. Найдены оценки нижних границ числа узлов равномерной сетки, необходимые для достижения назначенной точности сеточного решения задачи Коши для модельного уравнения Далквиста указанными методами.

Ключевые слова: конечно-разностные методы; глобальная погрешность; методы Рунге — Кутты; точность аппроксимации

Представлена в редакцию: 20.12.2019.

1. Введение

Большинство математических моделей в технике, физике, химической кинетике, экологии, метеорологии, представляют собой системы обыкновенных дифференциальных уравнений (ОДУ) вида

$$\dot{x} = f(t, x), \quad t \geq t_0, \tag{1}$$

где $x = (x^1, \ldots, x^m)^{\mathrm{T}}$, $f = (f^1(t,x), \ldots, f^m(t,x))^{\mathrm{T}}$ — вектор-столбцы искомой функции $x(t)$ и правой части $f(t,x)$ системы (1) размерности m ($1 \leq m \leq \infty$, $^{\mathrm{T}}$— знак транспонирования); $t_0 \in \mathrm{R}$ — некоторый начальный момент времени.

В большинстве случаев ставится задача Коши: найти решение $x(t)$ уравнения (1) с начальным значением

$$x(t_0) = x_0 = \Big(x^1(t_0), \ldots, x^m(t_0)\Big)^{\mathrm{T}}. \tag{2}$$

Для нахождения решений ОДУ применяются аналитические (точные), приближенно-аналитические и численные методы [1]–[5]. Получить точное решение означает выразить решение через элементарные, специальные функции или через интегралы от элементарных функций. Однако применить методы точного решения удается только для очень ограниченного круга уравнений.

Приближенно-аналитическими называются методы, позволяющие получить приближенное решение дифференциального уравнения в виде аналитического выражения. Среди этого класса следует выделить асимптотические методы, базирующиеся на разложении искомого решения системы (1) в формальный ряд по степеням некоторого малого параметра, с последующим усечением ряда и использовании его частных сумм в качестве приближенных решений (метод малого параметра [6], [7], возмущений [8], [9], методы Ляпунова, Пуанкаре [10], [11], методы осреднения Ван-дер-Поля [11], [12], Н.Н. Боголюбова, Ю.А. Митропольского [11], [13], и др.).

Численные методы предполагают получение приближенного значения решения ОДУ в виде таблицы приближенных значений искомой функции $x(t)$ для ряда значений независимой переменной t из интервала $[t_0,\ T]$, называемых сеткой разностного метода. Отметим, что в отличие от ранее указанных аналитических и приближенно-аналитических методов, решение задачи Коши (1), (2) в численных (конечно-разностных) методах ищется не на всем отрезке интегрирования $[t_0,\ T]$, а лишь на выбранной сетке.

В настоящей работе рассматриваются конечно-разностные методы — явный метод Эйлера первого порядка точности и метод Рунге — Кутты второго порядка точности — решения задачи Коши для модельного уравнения Далквиста. Приведены оценки погрешности указанных сеточных методов, найдены оценки нижних границ числа узлов равномерной сетки, необходимых для достижения назначенной точности сеточного решения задачи Коши.

2. Основные понятия теории конечно-разностных схем

Численные конечно-разностные методы позволяют получить значения искомого решения $x(t)$ задачи Коши

$$\dot{x} = f(t, x), \quad x(t_0) = x_0 \tag{3}$$

на отрезке $t \in [t_0,\ T]$ для данных вектор-столбцов начальных значений (2) и правой части (1).

Рассмотрим общий подход к вычислению приближенных решений задачи Коши (3) с помощью конечно-разностных методов (схем).

Разобьем отрезок $[t_0,\ T]$ точками $t_0 < t_1 < \ldots < t_n = T$. Набор этих точек в теории разностных схем называется сеткой, а сами точки — узлами сетки. Расстояние между соседними узлами называется шагом интегрирования $\tau = t_{k+1} - t_k$. Шаг может быть задан заранее (интегрирование с постоянным шагом) или может меняться в ходе вычислений. Для упрощения анализа в настоящей работе будем рассматривать равномерные сетки

$$\{t_k\}_{k=0}^{n} = \left\{ t_k = t_0 + k\tau,\ k = 0,\, 1,\, \ldots,\, n;\ \tau = \frac{T - t_0}{n} \right\}. \tag{4}$$

Под сеточным решением задачи Коши (3) понимается набор чисел $\{x_k\}_{k=0}^n$, где при каждом $k = 0,\, 1,\, \ldots,\, n$, x_k — значения точного решения конечно-разностного метода (схемы) в узлах сетки t_k. Сеточное решение используется для аппроксимации искомого точного решение дифференциальной задачи Коши, а точность конечно-разностного метода (схемы) определяет точность аппроксимации. Чтобы определить точность конечно-разностного метода (схемы) вводятся понятия локальной ошибки и глобальной погрешности. Локальной ошибкой конечно-разностного метода на k-м шаге называется разница между значениями сеточного решения x_k и решения $x(t)$ дифференциальной задачи Коши в узлах сетки, вычисленная в некоторой векторной норме в пространстве R^m:

$$\varepsilon_k = |x_k - x(t_k)|, \quad k = 1,\, 2,\, \ldots,\, n,$$

где $|\cdot|$ — выбранная норма в R^m; $x_k = (x_k^1,\, \ldots,\, x_k^m)^{\mathrm{T}}$ — значение сеточного решения задачи (3) в узле t_k; $x(t_k) = (x^1(t_k),\, \ldots,\, x^m(t_k))^{\mathrm{T}}$ — значение точного решения $x(t)$ дифференциальной задачи Коши в узле t_k (в том же узле сетки). Глобальную погрешность метода составляет совокупный вклад локальных ошибок по всем узлам сетки, определяемый сеточной функцией $\{\varepsilon_k\}_{k=1}^n$ по некоторой согласованной норме в пространстве сеточных функций:

$$\varepsilon(\tau) = \left\| \{\varepsilon_k\}_{k=1}^n \right\|.$$

Аппроксимация имеет порядок точности r, если существует такая, независящая от τ константа $M > 0$, что при $\tau \to 0$ глобальная погрешность конечно-разностного метода (схемы) допускает оценку

$$\varepsilon(\tau) \le M\tau^r. \tag{5}$$

Число r называется порядком точности конечно-разностного метода и его глобальная погрешность оценивается в неравенстве (5) значениями τ, M и r. Неравенство (5) имеет асимптотический характер и точное значение константы M заранее неизвестно, поэтому порядок точности конечно-разностного метода не зависит от выбора норм в определениях локальной ошибки и глобальной погрешности, и обычно, просто говорится, что для конечно-разностного метода порядка точности r его глобальная погрешность есть «O большое» от τ^r.

3. Постановка задачи и подход к решению

Пусть для некоторого заданного числа $\varepsilon > 0$ и конечно-разностного метода порядка точности r требуется обеспечить оценку глобальной погрешности в виде

$$\varepsilon(\tau) \le \varepsilon. \tag{6}$$

Из выражения (5) видим, что основным параметром, влияющим на оценку глобальной погрешности конечно-разностного метода является шаг интегрирования τ или, что то же самое, количество узлов сетки n.

З а м е ч а н и е 1. Строго говоря, формально при данном n количество узлов в сетке (4) равно $n+1$, но поскольку значение n определяет шаг интегрирования τ, который является ключевым параметром в характеристике конечно-разностных схем, будем отождествлять количество узлов в сетке с числом n «нестрого», исключая узел, отвечающий начальному моменту времени.

Определение 1. Конечно-разностный метод достигает назначенной точности $\varepsilon > 0$ на сетке (4), если его глобальная погрешность удовлетворяет условию (6).

Рассмотрим задачу достижения назначенной точности конечно-разностного метода при как можно меньшем количестве узлов сетки.

1. Достижение назначенной точности конечно-разностного метода. Решение задачи достижения назначенной точности ε конечно-разностной схемы непосредственно при помощи оценки (5), которое имеет вид

$$n = \left[\left(\frac{M}{\varepsilon} \right)^{1/r} \right] + 1$$

($[\,x\,]$ — целая часть числа x) при найденной константе M, может давать завышенный результат в силу специфичности характера оценки (5).

Действительно, например, в работе [14] приведена оценка погрешности метода Эйлера для задачи Коши (3) с правой частью $f(t,x)$, удовлетворяющей условию Липшица с константой L:

$$\varepsilon(\tau) \leq \frac{e^{TL}-1}{L}\, C\tau, \quad \tau = \frac{1}{n}, \tag{7}$$

где

$$C = \frac{1}{2} \max_{t_0 \leq t \leq T} |\ddot{x}(t)|$$

(считается, что решение задачи Коши является дважды непрерывно дифференцируемой функцией). Число узлов сетки (4), необходимое для достижения назначенной точности ε, можно получить из оценки (7) в виде

$$n = \left[\frac{e^{TL}-1}{\varepsilon L}\, C \right] + 1. \tag{8}$$

Далее в работе (см. 5) для модельной задачи Далквиста будет показано, что необходимое для достижения назначенной точности ε число узлов n в равномерной сетке существенно меньше, чем число n, найденное по формуле (8).

2. Явный анализ погрешности конечно-разностных методов. Более точно оценить необходимое для достижения назначенной точности ε число узлов n в равномерной сетке можно проводя явный анализ глобальной погрешности сеточного решения задачи Коши в каждом конкретном случае.

В настоящей работе аналитически исследуется погрешность явного метода Эйлера первого порядка точности для модельного уравнения Далквиста путем сравнения значений точных решений дифференциальной и разностной задач Коши в узлах равномерной сетки по модулю, а глобальная погрешность определяется максимумом модулей локальных погрешностей на выбранной сетке.

Оценка глобальной погрешности получается из неравенств, основанных на разложениях функций экспоненты и логарифма в ряды Тейлора и Меркатора и явно зависит от количества n узлов равномерной сетки.

Анализ полученной оценки глобальной погрешности показывает, что для достижения назначенной точности ε достаточно примерно в $1{,}7$ раза меньшее количество узлов равномерной сетки, чем число узлов n, найденное по формуле (8).

Далее проводится аналогичное исследование погрешности метода Рунге — Кутты второго порядка точности. Найденная оценка погрешности позволяет получить нижние границы для числа узлов равномерной сетки, на которой достигается назначенная точность сеточного решения, получаемого этим методом для модельного уравнения Далквиста.

4. Оценки погрешностей конечно-разностных схем первого и второго порядка точности на примере приближенного решения модельного уравнения Далквиста

В данном разделе будут приведены оценки глобальной погрешности явного метода Эйлера первого порядка точности

$$z_{k+1} = z_k + \tau f(t_k, z_k), \;\; k = 0,\, 1,\, \ldots,\, n-1, \quad z_0 = x_0, \tag{9}$$

и глобальной погрешности метода Рунге — Кутты второго порядка точности

$$\bar{z}_{k+1} = z_k + \tau f(t_k, z_k), \;\; k = 0,\, 1,\, \ldots,\, n-1, \quad z_0 = x_0; \tag{10}$$

$$z_{k+1} = z_k + \frac{\tau}{2}\Big(f(t_k, z_k) + f(t_{k+1}, \bar{z}_{k+1})\Big), \tag{11}$$

где $\bar{z}_{k+1}$ — некоторые вспомогательные векторы, задаваемые формулами (10), для приближенного решения задачи Коши модельного уравнения Далквиста

$$\dot{x} = ax, \quad m = 1, \quad x_0 = 1, \quad a = \mathrm{const} > 0, \quad t_0 = 0, \quad t \in [0,\ T] \tag{12}$$

на равномерной сетке

$$t_k = k\tau, \;\; k = 0,\, 1,\, \ldots,\, n, \quad \tau = \frac{T}{n}. \tag{13}$$

Как известно, точное решение модельной задачи Коши (12) задается формулой

$$x(t) = e^{at}, \quad 0 \le t \le T, \tag{14}$$

и, таким образом, глобальная погрешность конечно-разностной схемы на сетке (13) между точным решением модельной задачи Коши и его приближенным значением z_k определяется выражением

$$\varepsilon(\tau) = \max_{0\le k\le n} |e^{at_k} - z(t_k)| = \max_{0\le k\le n} |e^{a\tau k} - z_k|. \tag{15}$$

1. Оценка погрешности схемы Эйлера первого порядка точности. Оценим погрешность явного метода Эйлера (9) приближенного решения задачи Коши (12). В этом случае конечно-разностная схема метода Эйлера на равномерной сетке (13) имеет следующий вид:

$$\frac{z_{k+1}-z_k}{\tau} = az_k, \quad z_0 = 1, \ \ \tau = \frac{T}{n}, \ \ t_k = \frac{kT}{n}, \ \ k = 0,\, 1,\, \ldots,\, n-1. \tag{16}$$

Принимая во внимание (16) находим для любого $k = 0, 1, \ldots, n-1$

$$z_{k+1} = z_k + \tau a z_k = \left(1 + \frac{Ta}{n}\right) z_k,$$

откуда получаем, что при всех $k = 1, \ldots, n$

$$z_k = \left(1 + \frac{Ta}{n}\right)^k z_0 = \left(1 + \frac{Ta}{n}\right)^k. \tag{17}$$

Глобальная погрешность явной схемы Эйлера (9) на сетке (13) между точным решением модельной задачи Коши (12) и его приближенным значением z_k определяется выражением

$$\varepsilon(\tau) = \max_{0\le k\le n}\left| e^{\frac{aTk}{n}} - \left(1 + \frac{Ta}{n}\right)^k\right|. \tag{18}$$

Для оценки погрешности приближенного решения модельной задачи (12) методом Эйлера (9) справедливо следующее утверждение.

Теорема 1. Оценка погрешности (18) явного метода Эйлера (9) приближенного решения задачи Коши (12) имеет вид

$$\varepsilon(\tau) < \frac{e^{aT}T^2a^2}{2n}. \tag{19}$$

Доказательству теоремы предпошлем следующую лемму.

Лемма 1. Глобальная погрешность $\varepsilon(\tau)$ явного метода Эйлера (9) между точным решением $x(t) = e^{at}$, $t \in [0,\ T]$, и его приближенным значением z_k достигается в точке $t_n = T$ при $k = n$:

$$\varepsilon(\tau) = \max_{1\le k\le n} |x_k - z_k| = x(T) - z_n.$$

Доказательство. Покажем, что величина $\varepsilon(\tau)$ равна $x(T) - z_n$. В самом деле, согласно (17), для любого $k = 0,\, 1,\, \ldots,\, n-1$ имеем

$$|x_{k+1} - z_{k+1}| = \left| e^{\frac{Ta(k+1)}{n}} - \left(1 + \frac{Ta}{n}\right)^{k+1}\right| = \left| e^{\frac{Ta}{n}}\, e^{\frac{Tak}{n}} - \left(1 + \frac{Ta}{n}\right)\left(1 + \frac{Ta}{n}\right)^{k}\right|. \tag{20}$$

Неравенство $e^x > 1 + x$ при $x = \dfrac{Ta}{n}$ при любых натуральных n и k дает

$$\left(e^{\frac{Ta}{n}}\right)^k > \left(1 + \frac{Ta}{n}\right)^k. \tag{21}$$

Используя неравенство (21) заменим экспоненту e^{Ta} в (20) меньшей величиной, уберем модуль и получим для любого $k = 0,\, 1,\, \ldots,\, n-1$

$$|x_{k+1} - z_{k+1}| > \left(1 + \frac{Ta}{n}\right)\left(e^{\frac{Tak}{n}} - \left(1 + \frac{Ta}{n}\right)^k\right) = x_k - z_k,$$

откуда следует, что

$$\varepsilon(\tau) = \max_{1\le k\le n} |x_k - z_k| = \max_{1\le k\le n} (x_k - z_k) = x(T) - z_n.$$

Лемма 1 доказана.

Д о к а з а т е л ь с т в о теоремы (1). Покажем, что величина $\varepsilon(\tau)$ удовлетворяет неравенству (19):

$$\varepsilon(\tau) = x(T) - z_n < \frac{e^{aT}T^2a^2}{2n}.$$

Неравенство (30) с $x = \dfrac{Ta}{n}$ при любом натуральном n дает

$$\frac{Ta}{n} > \ln\Big(1 + \frac{Ta}{n}\Big) > \frac{Ta}{n} - \frac{T^2a^2}{2n^2},$$

откуда, умножая на n и применяя правило логарифма степени, получаем

$$Ta > \ln\Big(1 + \frac{Ta}{n}\Big)^n > Ta - \frac{T^2a^2}{2n}. \tag{22}$$

Экспонируем (22), переносим e^{Ta} влево и умножаем обе части неравенств на -1. В итоге получаем

$$0 < e^{Ta} - \Big(1 + \frac{Ta}{n}\Big)^n < e^{Ta} - e^{Ta - \frac{T^2a^2}{2n}},$$

или, вынося справа множитель e^{Ta} за скобку:

$$e^{Ta} - \Big(1 + \frac{Ta}{n}\Big)^n < e^{Ta}\Big(1 - e^{-\frac{T^2a^2}{2n}}\Big), \tag{23}$$

Из неравенства $e^x > 1 + x$ при $x = -T^2a^2/2n$ для любого натурального n получаем

$$e^{-\frac{T^2a^2}{2n}} > 1 - \frac{T^2a^2}{2n}. \tag{24}$$

Подставим (24) в (23) не нарушая неравенства:

$$e^{Ta} - \Big(1 + \frac{Ta}{n}\Big)^n < e^{Ta}\Big(1 - 1 + \frac{T^2a^2}{2n}\Big) = e^{Ta}\Big(\frac{T^2a^2}{2n}\Big),$$

откуда окончательно получаем искомую оценку:

$$\varepsilon(\tau) = x(T) - z_n = e^{Ta} - \Big(1 + \frac{Ta}{n}\Big)^n < \frac{e^{aT}T^2a^2}{2n}.$$

Теорема доказана.

Следствие 1. При $a = 1$, $T = 1$ погрешность метода Эйлера (9) равна $e/2n$:

$$\varepsilon(\tau) < \frac{e}{2n}. \tag{25}$$

2. Оценка погрешности схемы Рунге — Кутты второго порядка точности. Оценим погрешность метода Рунге — Кутты (11) второго порядка точности приближенного решения задачи Коши (12). Конечно-разностная схема метода Рунге — Кутты второго порядка точности на равномерной сетке (13) имеет следующий вид

$$z_{k+1} = z_k + \frac{\tau}{2}\Big[az_k + a(z_k + a\tau z_k)\Big] = \Big[1 + a\tau + 0{,}5a^2\tau^2\Big]z_k, \tag{26}$$

Из соотношений (26) при $z_0 = 1$ имеем

$$z_k = \left(1 + \frac{aT}{n} + \frac{a^2T^2}{2n^2}\right)^k, \quad k = 1,\, 2,\, \ldots,\, n. \tag{27}$$

Глобальная погрешность схемы Рунге — Кутты второго порядка точности (11) на сетке (13) между точным решением модельной задачи Коши (12) и его приближенным значением z_k определяется выражением

$$\varepsilon(\tau) = \max_{0\le k\le n}\left|e^{\frac{aTk}{n}} - \left(1 + \frac{aT}{n} + \frac{a^2T^2}{2n^2}\right)^k\right|. \tag{28}$$

Для оценки погрешности приближенного решения модельной задачи (12) методом Рунге — Кутты (11) справедливо следующее утверждение.

Теорема 2. Оцека погрешности (28) для метода Рунге — Кутты второго порядка точности задачи Коши (12) удовлетворяет следующему соотношению

$$\varepsilon(\tau) < e^{aT}\frac{a^2T^2}{2n^2}\left(1 + \frac{aT}{4n}\right). \tag{29}$$

Доказательству теоремы предпошлем две леммы.

Лемма 2. Для любого $x > 0$ верно неравенство

$$x > \ln(1 + x) > x - \frac{x^2}{2}. \tag{30}$$

Д о к а з а т е л ь с т в о. Неравенство слева в результате экспонирования переходит в известное неравенство $e^x > 1 + x$. Неравенство справа на интервале $(0,\ 1)$ вытекает из оценок для знакочередующегося ряда, сумма которого заключена между частичными суммами с четными и нечетными номерами. На промежутке $[0,\ +\infty)$ функция $\ln(1+x)$ возрастает, в то время как функция $x - x^2/2$ убывает. Поэтому очевидное неравенство $\ln 2 > 0{,}5$ при $x = 1$ сохраняется и при $x > 1$. Лемма доказана.

Лемма 3. Глобальная погрешность $\varepsilon(\tau)$ метода Рунге — Кутты второго порядка точности между точным решением $x(t) = e^{at}$, $t \in [0,\ T]$, и его приближенным значением z_k достигается в точке $t_n = T$ при $k = n$

$$\varepsilon(\tau) = \max_{1\le k\le n}|x_k - z_k| = x(T) - z_n.$$

Д о к а з а т е л ь с т в о. Покажем, что величина $\varepsilon(\tau)$ равна $x(T) - z_n$. В самом деле, согласно (27), для любого $k = 0, 1, \ldots, n-1$ имеем

$$|x_{k+1} - z_{k+1}| = \left| e^{\frac{Ta(k+1)}{n}} - \left(1 + \frac{aT}{n} + \frac{a^2T^2}{2n^2}\right)^{k+1} \right| =$$

$$= \left| e^{\frac{Ta}{n}} e^{\frac{Tak}{n}} - \left(1 + \frac{Ta}{n} + \frac{a^2T^2}{2n^2}\right)\left(1 + \frac{aT}{n} + \frac{a^2T^2}{2n^2}\right)^{k} \right|.$$

Рассматривая первые три члена в тейлоровском разложении функции e^x при $x = Ta/n$, убеждаемся в справедливости при любых натуральных n и k неравенства

$$\left(e^{\frac{Ta}{n}}\right)^k > \left(1 + \frac{aT}{n} + \frac{a^2T^2}{2n^2}\right)^{k+1},$$

используя которое аналогичным доказательству леммы 1 образом, получаем для любого $k = 0, 1, \ldots, n-1$ оценку

$$|x_{k+1} - z_{k+1}| = \left| e^{\frac{Ta(k+1)}{n}} - \left(1 + \frac{aT}{n} + \frac{a^2T^2}{2n^2}\right)^{k+1} \right| >$$

$$> \left(1 + \frac{Ta}{n} + \frac{a^2T^2}{2n^2}\right)\left(e^{\frac{Tak}{n}} - \left(1 + \frac{aT}{n} + \frac{a^2T^2}{2n^2}\right)^{k}\right) > x_k - z_k,$$

из которой следует, что

$$\varepsilon(\tau) = \max_{1 \le k \le n} |x_k - z_k| = \max_{1 \le k \le n} (x_k - z_k) = x(T) - z_n.$$

Лемма доказана.

Д о к а з а т е л ь с т в о теоремы 2. Покажем, что величина $\varepsilon(\tau)$ удовлетворяет неравенству (29):

$$\varepsilon(\tau) < e^{aT} \frac{a^2T^2}{2n^2}\left(1 + \frac{aT}{4n}\right).$$

Неравенство (30) при $x = a\tau + 0{,}5(a\tau)^2$ (напомним, что $\tau = T/n$ — шаг равномерной сетки (13)) для любого натурального n дает

$$\ln\left(1 + a\tau + 0{,}5(a\tau)^2\right) > a\tau(1 + 0{,}5a\tau) - 0{,}5(a\tau)^2(1 + 0{,}5a\tau)^2 =$$

$$= a\tau(1 + 0{,}5a\tau)\left(1 - 0{,}5a\tau(1 + 0{,}5a\tau)\right). \quad (31)$$

Подставим в первый сомножитель выражение $\tau = T/n$ (остальные τ оставляем, чтобы не загромождать формулы), и получим в (31) неравенство

$$\ln\left(1 + a\tau + 0{,}5(a\tau)^2\right) > \frac{T}{n} a(1 + 0{,}5a\tau)\left(1 - 0{,}5a\tau(1 + 0{,}5a\tau)\right),$$

умножая которое на n, получаем

$$n\ln\left(1 + a\tau + 0{,}5(a\tau)^2\right) > aT(1 + 0{,}5a\tau)\left(1 - 0{,}5a\tau(1 + 0{,}5a\tau)\right),$$

или, по свойству степени логарифма

$$\ln\Big(1 + a\tau + 0{,}5(a\tau)^2\Big)^n > aT(1 + 0{,}5a\tau)\Big(1 - 0{,}5a\tau(1 + 0{,}5a\tau)\Big).$$

Экспонируем последнее неравенство:

$$\Big(1 + a\tau + 0{,}5(a\tau)^2\Big)^n > e^{aT(1+0{,}5a\tau)(1-0{,}5a\tau(1+0{,}5a\tau))}.$$

Вычтем обе части последнего неравенства из e^{aT}. При этом знак неравенства изменится на противоположный. С учетом (27) получим

$$e^{aT} - z_n < e^{aT} - e^{aT(1+0{,}5a\tau)(1-0{,}5a\tau(1+0{,}5a\tau))}. \tag{32}$$

Преобразуем выражение в показателе экспоненты:

$$aT(1 + 0{,}5a\tau)(1 - 0{,}5a\tau(1 + 0{,}5\,a\tau)) = aT(1 + 0{,}5a\tau)(1 - 0{,}5a\tau - 0{,}25\,a^2\tau^2) =$$
$$= aT(1 - 0{,}5a^2\tau^2 - 0{,}125a^3\tau^3) = aT(1 - 0{,}5a^2\tau^2(1 + 0{,}25a\tau)).$$

Возвращаясь к неравенству (32), видим, что

$$e^{aT} - z_n < e^{aT} - e^{aT(1-0{,}5a^2\tau^2(1+0{,}25a\tau))} = e^{aT}\Big(1 - e^{-0{,}5a^2\tau^2(1+0{,}25a\tau)}\Big).$$

Неравенство $e^x > 1 + x$ при $x = -0{,}5a^2\tau^2(1 + 0{,}25a\tau)$ для любом натурального n дает

$$e^{-0{,}5a^2\tau^2(1+0{,}25a\tau)} > 1 - 0{,}5a^2\tau^2(1 + 0{,}25a\tau),$$

откуда получаем

$$e^{aT} - z_n < e^{Ta}\Big(1 - 1 + 0{,}5a^2\tau^2(1 + 0{,}25a\tau)\Big) = e^{Ta} \cdot 0{,}5a^2\tau^2(1 + 0{,}25a\tau).$$

Применяя результат Леммы 3 и вспоминая, что $\tau = T/n$, окончательно получаем искомую оценку:

$$\varepsilon(\tau) = x(T) - z_n < e^{Ta}\,\frac{a^2T^2}{2n^2}\Big(1 + \frac{aT}{4n}\Big).$$

Теорема 2 доказана.

Следствие 2. При $a = 1, T = 1$ погрешность метода Рунге — Кутты (11) второго порядка точности равна:

$$\varepsilon(\tau) < \frac{e}{2n^2}\Big(1 + \frac{1}{4n}\Big). \tag{33}$$

В следующем разделе найденные оценки используются для получения нижних границ числа узлов равномерной сетки рассматриваемых методов Эйлера и Рунге — Кутты решения задачи Коши модельного уравнения Далквиста (12) для достижения назначенных точностей сеточного решения $\varepsilon = 0{,}05;\ 0{,}01;\ 0{,}005$.

5. Примеры

В данном разделе будем предполагать, что в модельном уравнении Далквиста (12) величины a и T равны единице. Оценим число узлов сетки, необходимое для достижения точностей $\varepsilon_1 = 5 \cdot 10^{-2}$, $\varepsilon_2 = 10^{-2}$, $\varepsilon_3 = 5 \cdot 10^{-3}$.

Оценка числа узлов равномерной сетки для метода Эйлера. Оценка глобальной погрешности явного метода Эйлера приближенного решения задачи Коши при $a = 1, T = 1$ согласно следствию 1 определяется формулой (25)

$$\varepsilon(\tau) < \frac{e}{2n}.$$

Из этой формулы заключаем, что наименьшее число узлов равномерной сетки для достижения назначенной точности ε сеточного решения задачи Коши методом Эйлера находится из условия

$$n(\varepsilon) = \left[\frac{e}{2\varepsilon}\right] + 1.$$

В частности, для $\varepsilon_1 = 5 \cdot 10^{-2}$ получаем

$$n(\varepsilon_1) = \left[\frac{e}{2 \cdot 5 \cdot 10^{-2}}\right] + 1 = [27{,}1828] + 1 = 28.$$

Аналогично находим для $\varepsilon_2 = 10^{-2}$

$$n(\varepsilon_2) = \left[\frac{e}{2 \cdot 10^{-2}}\right] + 1 = [135{,}914] + 1 = 136,$$

и для $\varepsilon_3 = 5 \cdot 10^{-3}$

$$n(\varepsilon_3) = \left[\frac{e}{2 \cdot 5 \cdot 10^{-3}}\right] + 1 = [271{,}8281] + 1 = 272.$$

Сравнение оценок числа узлов равномерной сетки. Сравним полученные значения $n(\varepsilon_1)$, $n(\varepsilon_2)$ и $n(\varepsilon_3)$ с оценочными значениями числа узлов m, необходимых для достижения назначенной точности, которые могли бы быть найдены по формуле (8). Несложно проверить, что при $a = 1$ и $T = 1$ значение константы L в выражении (8) также равно единице, а константа $C = e/2$. Подставляя эти значения в (8) получаем

$$m = \left[\frac{e\,(e-1)}{2\varepsilon}\right] + 1, \tag{34}$$

и для ε_1, ε_2 и ε_3 находим соответственно $m(\varepsilon_1) = 47$, $m(\varepsilon_2) = 234$, $m(\varepsilon_3) = 468$. Таким образом, видим, что отношения $m(\varepsilon_i)/n(\varepsilon_i)$ равны примерно $1,7$ ($47/28 \approx 1,67857$, $234/136 \approx 1,72058$, $468/272 \approx 1,72058$).

З а м е ч а н и е 2. Константа M из (5) в оценках (7) и (25) имеет вид $M = e(e-1)/2$ и $M = e/2$ соответственно, поэтому соотношение 1,7 вытекает из отношения значений этих констант, равного $\dfrac{e(e-1)}{e} = e - 1 \approx 1{,}7$.

Оценка числа узлов равномерной сетки для метода Рунге — Кутты. Оценка глобальной погрешности явного метода Рунге — Кутты второго порядка точности приближенного решения задачи Коши при $a = 1$, $T = 1$ согласно следствию 2 определяется формулой (33)

$$\varepsilon(\tau) < \frac{e}{2n^2}\Big(1 + \frac{1}{4n}\Big).$$

Из этой формулы заключаем, что наименьшее число узлов равномерной сетки для достижения заданной точности ε приближенного решения задачи Коши методом Рунге — Кутты второго порядка точности находится из условия

$$\frac{4n^3}{4n+1} > \frac{e}{2\varepsilon}.$$

Данное условие можно упростить. Разделим $4n^3$ на $4n + 1$ с остатком и получим

$$\frac{4n^3}{4n+1} = n^2 - \frac{1}{4}\,n + \frac{1}{16} - \frac{1}{16(4n+1)},$$

откуда видим, что при любом натуральном $n \geq 1$ справедливы неравенства

$$n^2 - \frac{1}{4}\,n < \frac{4n^3}{4n+1} < n^2.$$

Воспользуемся нижней оценкой $n^2 - \dfrac{1}{4}n < \dfrac{4n^3}{4n+1}$. Решив квадратное неравенство

$$x^2 - \frac{x}{4} - \frac{e}{2\varepsilon} > 0,$$

получим следующее условие нахождения наименьшего числа узлов равномерной сетки для достижения заданной точности ε приближенного решения задачи Коши методом Рунге — Кутты второго порядка точности:

$$n(\varepsilon) = \left[\frac{1}{8} + \sqrt{\frac{1}{64} + \frac{e}{2\varepsilon}}\,\right] + 1.$$

В частности, для $\varepsilon_1 = 5 \cdot 10^{-2}$ получаем

$$n(\varepsilon_1) = \left[\frac{1}{8} + \sqrt{\frac{1}{64} + \frac{e}{2 \cdot 5 \cdot 10^{-2}}}\,\right] + 1 = [5{,}34021] + 1 = 6.$$

Аналогично находим для $\varepsilon_2 = 10^{-2}$

$$n(\varepsilon_2) = \left[\frac{1}{8} + \sqrt{\frac{1}{64} + \frac{e}{2 \cdot 10^{-2}}}\,\right] + 1 = [11{,}78389] + 1 = 12,$$

и для $\varepsilon_3 = 5 \cdot 10^{-3}$

$$n(\varepsilon_3) = \left[\frac{1}{8} + \sqrt{\frac{1}{64} + \frac{e}{2 \cdot 5 \cdot 10^{-3}}}\,\right] + 1 = [16{,}61268] + 1 = 17.$$

Заключение

Рассмотрена задача достижения назначенной точности с наименьшим количеством узлов сетки для прямых конечно-разностных схем Рунге — Кутты первого и второго порядка точности. Для оценки необходимого числа узлов в равномерной сетке проведен явный анализ глобальной погрешности сеточного решения задачи Коши для каждого из названных методов. Аналитически исследованы погрешности явного метода Эйлера первого порядка точности и метода Рунге — Кутты второго порядка точности для модельного уравнения Далквиста путем сравнения значений точных решений дифференциальной и разностной задач Коши в узлах равномерной сетки. Найдены оценки нижних границ числа узлов равномерной сетки, необходимые для достижения назначенной точности сеточного решения задачи Коши указанными методами.

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований, грант 19-07-00686 а.

Список литературы

1. Рябенький В.С. Введение в вычислительную математику: учеб. пособие. 3-е изд. М.: Физматлит, 2008. 284 с.
2. Тихонов А.Н., Васильева А.Б., Свешников А.Г. Дифференциальные уравнения: учебник. 4-е изд. М.: Физматлит, 2005. 253 с.
3. Бахвалов Н.С., Жидков Н.П., Кобельков Г.М. Численные методы: учеб. пособие. 4-е изд. М.: БИНОМ. Лаборатория знаний, 2006. 636 с.
4. Лобанов А.И., Петров И.Б. Математическое моделирование нелинейных процессов. М.: Юрайт, 2019. 255 с.
5. Эльсгольц Л.Э. Качественные методы в математическом анализе. 3-е изд. М.: КомКнига: УРСС, 2010. 304 с.
6. Федорюк М.В. Асимптотические методы для линейных обыкновенных дифференциальных уравнений. 3-е изд. М.: Либроком, 2015. 352 с.
7. Мищенко Е.Ф., Розов Н.Х. Дифференциальные уравнения с малым параметром и релаксационные колебания. М.: Наука, 1975. 248 с.
8. Найфэ А.Х. Введение в теорию возмущений: учебник: пер. с англ. М.: Мир, 1984. 535 с. [Nayfeh A.H. Introduction to perturbation techniques. N.Y.: Wiley, 1981. 519 p.].
9. Мышкис А.Д. Элементы теории математических моделей. 2-е изд. М.: Эдиториал УРСС, 2004. 192 с.
10. Малкин И.Г. Методы Ляпунова и Пуанкаре в теории нелинейных колебаний. 3-е изд. М.: Эдиториал УРСС, 2010. 248 с.

11. Моисеев Н.Н. Асимптотические методы нелинейной механики: учеб. пособие. М.: Наука, 1969. 380 с.
12. Лобанов А.И., Петров И.Б. Вычислительные методы для анализа моделей сложных динамических систем: учеб. пособие. Ч. 1. М.: МФТИ, 2000. 168 с.
13. Боголюбов Н.Н., Митропольский Ю.А. Асимптотические методы в теории нелинейных колебаний. 3-е изд. М.: Физматгиз, 1963. 410 с.
14. Hairer E., Lubich Ch. Numerical solution of ordinary differential equations // The Princeton companion to applied mathematics / Ed. by N.J. Higham. Princeton; Oxf.: Princeton Univ. Press, 2015. Pp. 293–305.

Mathematics and Mathematical Modeling, 2019, no. 5, pp. 32–48.

DOI: **10.24108/mathm.0519.0000205**

Mathematics & Mathematical Modelling

Electronic journal

http://mathmelpub.ru ISSN 2412-5911

Error Margin Analysis of Certain Explicit Finite-difference Methods to Solve the Cauchy Problem for Dahlquist Model Equation

A. A. Akhrem[1], A. P. Nosov[1,*], V. Z. Rakhmankulov[1], K. V. Yuzhanin[1]

[1]Federal Research Center "Informatics and Management" of the Russian Academy of Sciences, Moscow, Russia

[*] nosov@isa.ru

Keywords: finite-different methods, global margin of error, Runge — Kutta methods, approximation accuracy

Received: 20.12.2019.

The aim of the paper is to estimate the minimum appropriate number of nodes on a uniform grid (maximum integration step) to obtain a given accuracy for the finite-difference Runge — Kutta methods of the first and second orders of accuracy for the Dahlquist model equation.

The error of finite-difference methods is analytically investigated by explicit comparing the values of the exact solutions of the differential and difference Cauchy problems in the nodes of a uniform grid in modulus, and the global error is determined by the maximum of the modules of the local errors on the selected grid. The estimates of the global error are obtained from the inequalities based on the expansions of the functions of the exponent and the logarithm in the Taylor and Mercator series, and clearly depend on the number of nodes of the uniform grid.

The bottom of the number of nodes of the uniform grid that is required to have the desirable accuracy to solve the Cauchy problem by above methods is obtained.

The obtained estimate of the global error of the direct Euler method for the Dahlquist model equation substantially refines the similar estimate from the paper (Hairer E., and Lubich C. Numerical Solution of Ordinary Differential Equations) and enables us to use an integration step of 1.7 times more in value, keeping the given approximation accuracy.

The accuracy order of the finite-difference schemes in the theory of numerical methods for integrating differential equations provides a relationship between the global error of the method and the integration step, however, it does not allow us to directly express the approximation accuracy on the given grid, and therefore, an optimal integration step is most often determined

experimentally. The paper studies such a relationship explicitly as a model example and shows one of the possible ways to obtain analytical estimates of the integration step for a given approximation accuracy.

A direct study of the global error of finite-difference schemes is important in problems where a trade-off between the approximation accuracy and the complexity (amount of computation) is of importance when the number of grid nodes matters. In this regard, it is of interest to extend similar studies of error estimation to the other finite-difference schemes, namely Runge — Kutta methods of higher orders of accuracy and multistep methods.

The results obtained can be useful for solving the tasks of computer modeling and computer-based learning.

References

1. Riaben'kij V.S. *Vvedenie v vychislitel'nuyu matematiku* [Introduction to computational mathematics]: a textbook. 3rd ed. Moscow: Fizmatlit Publ., 2008. 284 p. (in Russian).
2. Tikhonov A.N., Vasil'eva A.B., Sveshnikov A.G. *Differentsial'nye uravneniia* [Differential equations]: a textbook. 4th ed. Moscow: Fizmatlit Publ., 2005. 253 p. (in Russian).
3. Bakhvalov N.S., Zhidkov N.P., Kobel'kov G.M. *Chislennye metody* [Numerical methods]: a textbook. 4th ed. Moscow: BINOM. Laboratoriia znanij Publ., 2006. 636 p. (in Russian).
4. Lobanov A.I., Petrov I.B. *Matematicheskoe modelirovanie nelinejnykh protsessov* [Mathematical modeling of nonlinear processes]. Moscow: Urajt Publ., 2019. 255 p. (in Russian).
5. El'sgol'ts L.E. *Kachestvennye metody v matematicheskom analize* [Qualitative methods in mathematical analysis]. 3rd ed. Moscow: KomKniga Publ.: URSS Publ., 2010. 304 p. (in Russian).
6. Fedoryuk M.V. *Asimptoticheskie metody dlia linejnykh obyknovennykh differentsial'nykh uravnenij* [Asymptotic methods for linear ordinary differential equations]. 3rd ed. Moscow: Librokom Publ., 2015. 352 p. (in Russian).
7. Mishchenko E.F., Rozov N.Kh. *Differentsial'nye uravneniia s malym parametrom i relaksatsionnye kolebaniia* [Differential equations with small parameter and relaxation oscillations]. Moscow: Nauka Publ., 1975. 248 p. (in Russian).
8. Nayfeh A.H. *Introduction to perturbation techniques*. N.Y.: Wiley, 1981. 519 p. (Russ. ed.: Nayfeh A.H. *Vvedenie v teoriyu vozmushchenij*. Moscow: Mir Publ., 1984. 535 p.).
9. Mushkis A.D. *Elementy teorii matematicheskikh modelej* [Elements of the theory of mathematical models]. 2nd ed. Moscow: Editorial URSS Publ., 2004. 192 p. (in Russian).
10. Malkin I.G. *Metody Liapunova i Puankare v teorii nelinejnykh kolebanij* [Liapunov and Poincare methods in the theory of nonlinear oscillations]. 3rd ed. Moscow: Editorial URSS Publ., 2010. 248 p. (in Russian).

11. Moiseev N.N. *Asimptoticheskie metody nelinejnoj mekhaniki* [Asymptotic methods of nonlinear mechanics]: a textbook. Moscow: Nauka Publ., 1969. 380 p. (in Russian).

12. Lobanov A.I., Petrov I.B. *Vychislitel'nye metody dlia analiza modelej slozhnykh dinamicheskikh system* [Computational methods for analyzing models of complex dynamic systems]: a textbook. Pt. 1. Moscow: MPhTI Publ., 2000. 168 p. (in Russian).

13. Bogolyubov N.N., Mitropol'skij Yu.A. *Asimptoticheskie metody v teorii nelinejnykh kolebanij* [Asymptotic methods in the theory of nonlinear oscillations]. 3rd ed. Moscow: Fizmatgiz Publ., 1963. 410 p. (in Russian).

14. Hairer E., and Lubich Ch. Numerical solution of ordinary differential equations. *The Princeton companion to applied mathematics* / Ed. by N.J. Higham. Princeton; Oxf.: Princeton Univ. Press, 2015. Pp. 293–305.