

ФЕДЕРАЛЬНОЕ АГЕНТСТВО ПО ОБРАЗОВАНИЮ

Государственное образовательное учреждение высшего профессионального образования
«Уральский государственный университет им. А.М. Горького»

ИОНЦ «Физика в биологии и медицине»

Математико-механический факультет

Кафедра вычислительной математики

Функционально-дифференциальные уравнения в биологии и медицине

Учебное пособие

Подпись руководителя ИОНЦ

Дата

Екатеринбург
2008

В.Г.Пименов

**Функционально-дифференциальные
уравнения
в биологии и медицине**

Екатеринбург, 2008

Оглавление

1. Типы дифференциальных уравнений с запаздыванием	4
1.1. Некоторые типы ФДУ	4
2. Модели в биологии и медицине, описываемые ФДУ	7
2.1. Модель Хатчинсона	7
2.2. Модель Лотке-Вольтерры	9
2.3. Модели эпидемий	10
2.4. Модели в иммунологии	11
2.5. Модель Марчука	11
2.6. LCMV-модель	13
2.7. Некоторые другие модели	14
3. Одношаговые численные методы для ФДУ	16
3.1. Основные предположения	16
3.2. Численный метод Эйлера с кусочно-постоянной интерполяцией	17
3.3. Способы интерполяции и экстраполяции предыстории дискретной модели	19
3.4. Явные методы типа Рунге–Кутты	24
3.5. Порядок невязки ЯРК-методов	27
3.6. Неявные методы типа Рунге–Кутты	30
4. Многошаговые и некоторые другие численные методы для ФДУ	36
4.1. Многошаговые методы	36
4.2. Многошаговые методы, не требующие разгона	38
4.3. Методы Нордсика	39
4.4. Методы, использующие вычисление старших производных	41

4.5. Другие методы, основанные на разделении конечномерной и бес- конечномерной фазовых составляющих	44
5. Общая схема численных методов	47
5.1. Введение	47
5.2. Дискретная модель и порядок сходимости	48
5.3. Методика классификации численных моделей ФДУ	53
5.4. Необходимые и достаточные условия сходимости с порядком p	58
5.5. Асимптотическое разложение глобальной погрешности	61
6. Алгоритмы с переменным шагом и некоторые вопросы компьютер- ной реализации численных моделей	68
6.1. ЯРК-методы с переменным шагом	68
6.2. Способы интерполяции и экстраполяции расширенной предысто- рии дискретной модели	72
6.2.1. Выбор длины шага	75
6.3. Тестовые задачи	79
Библиография	88

Глава 1.

Типы дифференциальных уравнений с запаздыванием

Динамические модели, применяемые в биологии и медицине, зачастую должны учитывать зависимость будущего развития процесса не только от настоящего состояния, но от предыстории развития процесса. Математическим описанием таких моделей служат дифференциальные уравнения с запаздыванием, называемые также уравнениями с последствием или функционально-дифференциальными уравнениями (сокращенно ФДУ), которые являются обобщением обыкновенных дифференциальных уравнений (ОДУ). Система ОДУ записывается в виде

$$\dot{x} = f(t, x(t)),$$

где t – независимая переменная (трактуемая обычно как время), $x(t)$ – искомая векторная функция, x называется фазовым вектором.

Система ФДУ записывается в виде

$$\dot{x} = f(t, x(t), x_t(\cdot)),$$

где t – независимая переменная, x – фазовый вектор, $x_t(\cdot) = \{x(t+s), -\tau \leq s < 0\}$ – функция-предыстория фазового вектора, действующая на систему.

Рассмотрим несколько простых примеров.

1.1. Некоторые типы ФДУ

Пример 1.1. Рассмотрим уравнение

$$\dot{x}(t) = x(t - \tau), \quad (1.1)$$

($t \geq 0$, $\tau = 1$, x – одномерный фазовый вектор) с начальным условием $x(0) \equiv 1$. Для определения единственного решения уравнения (1.1) при $t > 0$ недостаточно знания начального фазового вектора $x(0)$, необходимо задать начальную функцию-предысторию $y^0(\cdot)$. Пусть, например, $y^0(s) = 1$ при $-1 \leq s < 0$. Тогда:

а) если $0 \leq t \leq 1$, то известна функция $x(t-1) = y^0(t-1) = 1$, и, интегрируя исходное уравнение, получаем

$$x(t) = x(0) + \int_0^t x(t-1)dt = 1 + t,$$

т.е. на этом участке решение — линейная функция;

б) если $1 \leq t \leq 2$, то $x(t-1) = 1 + (t-1) = t$, и, интегрируя исходное уравнение, получаем

$$x(t) = x(1) + \int_1^t x(t-1)dt = 2 + \int_1^t tdt = \frac{t^2}{2} + \frac{3}{2},$$

т.е. на этом участке решение — парабола.

На последующих участках решение также можно выписать, при этом его гладкость повышается с течением времени. Применённый метод интегрирования уравнений с запаздыванием называется *методом шагов*.

Отметим, что в общем случае решение на интервале $[0, \infty)$ более общего уравнения

$$\dot{x}(t) = a x(t - \tau) \quad (1.2)$$

($a > 0$) с постоянной начальной функцией

$$y^0(s) = c, \quad -\tau \leq s \leq 0 \quad (1.3)$$

($c \in \mathbb{R}$) имеет вид

$$x(t) = c \sum_{m=0}^{\left[\frac{t}{\tau}\right]+1} \frac{a^m \left(t - (m-1)\tau\right)^m}{m!}, \quad (1.4)$$

где $[\zeta]$ обозначает целую часть числа $\zeta \in \mathbb{R}$. \square

Рассмотренный пример относится к типу *уравнений с постоянным сосредоточенным запаздыванием*

$$\dot{x}(t) = f(t, x(t), x(t - \tau)).$$

Если правая часть этого уравнения от $x(t)$ не зависит, то запаздывание называется *чистым*:

$$\dot{x}(t) = f(t, x(t - \tau)).$$

Величина запаздывания может быть переменной.

Пример 1.2. Рассмотрим уравнение

$$\dot{x}(t) = x(t - (e^{-t} + 1)) + \cos t - \sin(t - e^{-t} - 1),$$

где $t \geq 0$, x — одномерный фазовый вектор. Величина запаздывания $\tau(t) = e^{-t} + 1$, причем $1 < \tau(t) \leq 2$.

Если начальные условия задаются функцией $y^0(t) = \sin t$ при $-2 \leq t \leq 0$, то непосредственной подстановкой проверяется, что функция $x(t) = \sin t$ при $t \geq 0$ является решением, соответствующим данной начальной функции. \square

Этот пример относится к типу ФДУ с *переменным сосредоточенным запаздыванием*

$$\dot{x}(t) = f(t, x(t), x(t - \tau(t))) .$$

Уравнение с *несколькими сосредоточенными запаздываниями* имеет вид

$$\dot{x}(t) = f(t, x(t - \tau_1), x(t - \tau_2), \dots, x(t - \tau_k)) .$$

Запаздывание может быть *распределённым*.

Пример 1.3. Рассмотрим уравнение

$$\dot{x}(t) = \int_{-1}^0 x_t(s) ds = \int_{t-\tau}^t x(\xi) d\xi . \quad (1.5)$$

Продифференцировав обе части этого уравнения, получим

$$\ddot{x}(t) = x(t) - x(t - \tau) . \quad (1.6)$$

Будем искать решение этого уравнения в виде $x = e^{\lambda t}$, в результате подстановки получаем трансцендентное характеристическое уравнение

$$\lambda^2 - 1 = -e^{-\lambda\tau} ,$$

которое имеет ненулевой вещественный корень λ_0 . Этот корень можно найти численно: $\lambda_0 \approx 0.7$ при $\tau = 1$.

Таким образом, если начальные условия задать в виде $x(s) = e^{\lambda_0 s}$, $s \in [-1, 0]$, то при $t \geq 0$ решение уравнения (1.5) имеет вид $x(t) = e^{\lambda_0 t}$. \square

Отметим, что уравнение (1.6) не эквивалентно уравнению (1.5), так как имеет дополнительные решения, например, $x(t) \equiv 1$.

Этот пример относится к типу уравнений с *распределённым запаздыванием в виде интеграла*

$$\dot{x}(t) = f(t, x(t), \int_{-\tau}^0 \alpha(t, s, x(t+s)) ds) .$$

Интегродифференциальные уравнения Вольтерра

$$\dot{x}(t) = f(t, x(t), \int_0^t \alpha(t, s, x(s)) ds) ,$$

сводятся заменой $s = s - t$ к виду

$$\dot{x}(t) = f(t, x(t), \int_{-t}^0 \alpha(t, s - t, x(s - t)) ds) .$$

Существуют и более сложные виды запаздываний. Все эти уравнения можно объединить формулой

$$\begin{aligned} \dot{x} &= f(t, x(t), x_t(\cdot)) , \\ x_t(\cdot) &= \{x(t+s), -\tau \leq s < 0\} . \end{aligned}$$

Приведённые выше Примеры 1.1 — 1.3 являются хорошими тестами для проверки численных методов, так как в них явно выписывается точное решение. В большинстве же ФДУ указанных типов точное решение выписать невозможно, поэтому особенно актуальна задача разработки приближенных методов нахождения решения.

Следует отметить, что приведенная классификация ФДУ не является полной, а скорее приводит ряд примеров.

Далее рассматриваются только случаи, когда величина запаздывания неотрицательна: $0 \leq \tau(t)$. Если величина запаздывания $\tau(t) < 0$, то такие уравнения называются *уравнениями опережающего типа*, в них поведение решения может зависеть от будущего.

Уравнением *нейтрального типа* называется уравнение вида

$$\dot{x}(t) = f(t, x(t), x(t - \tau(t)), \dot{x}(t - \tau(t))) .$$

Такие уравнения также далее не рассматриваются.

Глава 2.

Модели в биологии и медицине, описываемые ФДУ

Введение запаздывания в дифференциальные уравнения, описывающие какой-либо биологический процесс, является естественным математическим приемом. Во многих задачах запаздывание имеет конкретный смысл. Например, в задаче о популяциях, условно классифицируемых как "хищник" и "жертва", оно может учитывать возраст части популяции или другие характеристики их развития, рождаемости или вымирания. Во многих моделях запаздывание вводится как характеристика малоизученных процессов, которая на данном этапе построения модели в неё не включаются. Это может быть, например, время транспорта молекул от места их синтеза к месту их включения в систему реакций; время формирования клеток определённого типа, участвующих в иммунной реакции; длительность реакции части популяции на лимитирующие факторы окружающей среды.

2.1. Модель Хатчинсона

Рассмотрим сначала две дифференциальные модели, не содержащие запаздывания.

Модель экспоненциального роста

Модель экспоненциального роста (убывания) является простейшей из дифференциальных моделей и широко применяется в таких науках, как экономика (закон прибавочной стоимости), физика (явление радиоактивного распада), биология (теория популяций), история (рост народонаселения). Рассмотрим модель в терминах теории популяций. Пусть время t – независимая переменная, количество особей в популяции N – искомая функция. Если количество особей велико, то можно считать $N(t)$ непрерывной функцией. Из многочисленных наблюдений установлено, что скорость роста количества особей пропорциональна количеству особей. Коэффициент пропорциональности α равен количеству рождений за единицу времени минус количество смертей за это же время и может считаться постоянным. Получаем дифференциальное уравнение первого порядка

$$\dot{N} = \alpha N. \quad (2.1)$$

Решение этого уравнения – экспонента $N = N_0 \exp \alpha t$, причем количество особей растёт до бесконечности. Эта модель описывает явление лишь в первом приближении, не учитывает многих факторов, главным из которых является неограниченность экологической ниши.

Логистическое уравнение

Более точная модель имеет вид

$$\dot{N} = \alpha N - \beta N^2 \quad (2.2)$$

и называется логистическим уравнением. Последнее слагаемое учитывает влияние фактор борьбы за пищу каждой особи с каждой. Это уравнение имеет два положения равновесия $N = 0$ и $\bar{N} = \frac{\alpha}{\beta}$. Если задано начальное число особей $0 < N_0 < \bar{N}$, то решение (логистическая кривая) асимптотически приближается к \bar{N} . Путем нормировки $x = \frac{N}{\bar{N}}$, это уравнение приводится к виду, содержащему один параметр:

$$\dot{x} = \alpha x(1 - x). \quad (2.3)$$

Модель Хатчинсона и её обобщения

Одной из первых моделей в биологии, в которых было введено запаздывание для объяснения проявления периодичности, была модель Хатчинсона [39].

$$\dot{x} = r(1 - \frac{x(t - \tau)}{K})x, \quad (2.4)$$

где r – мальтусов коэффициент линейного роста, τ – например, возраст производителей вида, K – средняя численность популяции, определяемая емкостью среды.

При $r\tau > \pi/2$ уравнение имеет в качестве решения периодические колебания. Этот результат не так просто получить аналитически. Поясним качественную картину, взяв для простоты $r = K = 1$, а также считая, что в начальный момент $x(0) < 1$. Предположим, что в некоторый момент (первый) t_1 будет $x(t_1) = 1$, тогда считая, что $\tau < t_1$, поскольку $x(t_1 - \tau) < 1$, получаем $\dot{x}(t_1) > 0$. Рост решения будет продолжаться, пока $x(t - \tau) < 1$. После $t_2 = t_1 + \tau$ начнется убывание $x(t)$ и если π не слишком мало, то начиная с некоторого момента t_3 будет вновь $x(t) < 1$. Возможность колебательного решения, таким образом, правдоподобна, что подтверждается численными расчетами.

Этот пример показывает заметный эффект дестабилизации, который может вызвать запаздывание на фоне стабилизирующего влияния лимитированных ресурсов. С помощью простой модели (2.4) были предприняты удачные попытки объяснения различных случаев циклического изменения численности популяции. Так, Мей [47] получил качественное совпадение результата теоретического исследования с экспериментальными данными для зеленой падальной мухи, в других работах модель с запаздыванием позволила объяснить 4-летний цикл популяции пашенной полевки. Различные примеры влияния возрастной структуры популяции на динамику ее численности приведены в [20, 21].

Во многих моделях рассматриваются обобщения уравнения Хатчинсона. Так Ю.С. Колесов показал, что уравнение (2.4) достаточно хорошо описывает динамику изменения численности многих массовых видов животных, так как у таких видов средняя продолжительность жизни обычно лишь немногим больше τ . Если же это условие нарушается, то можно перейти к уравнению

$$\dot{x} = r(1 - \frac{\alpha_1 x(t - \tau_1) + \alpha_2 x(t - \tau_2)}{K})x, \quad (2.5)$$

где параметры α_1 и α_2 характеризуют вклад соответствующей возрастной группы в рост популяций. Оказывается, что если $\tau_1 = 1$, а $\tau_2 = 2$, то некоторый процент долгожителей приводит к уменьшению осцилляции численности. Если же $\tau_1 = 2$, а $\tau_2 = 3$, то наличие любого количества долгожителей приводит к увеличению осцилляции численности. Отсюда, в частности, следует объяснение причины более интенсивных колебаний численности зайца-беляка в Якутии по сравнению с ондатрой.

Другой путь обобщения уравнения Хатчинсона – введение распределенного запаздывания, например, путем перехода к уравнению

$$\dot{x} = rx(1 - \frac{1}{K} \int_0^\tau x(t-s)ds), \quad (2.6)$$

или к другим модификациям и усложнениям модели (2.4).

2.2. Модель Лотке-Вольтерры

Основной задачей математической экологии является описание взаимодействующих популяций, и здесь основополагающей является модель Лотке-Вольтерры [7]. В исходной постановке эта система с распределенным запаздыванием

$$\begin{aligned} \dot{x}(t) &= [\varepsilon_1 - \gamma_1 y(t) - \int_{-\tau}^0 F_1(s)y(t+s)ds]x(t), \\ \dot{y}(t) &= [\gamma_2 x(t) - \varepsilon_2 + \int_{-\tau}^0 F_1(s)x(t+s)ds]y(t), \end{aligned} \quad (2.7)$$

где $x(t)$ и $y(t)$ суть численности (плотности популяций) жертв и хищников соответственно.

В большинстве ранних моделей эффект запаздывания игнорировался, и полученная система обыкновенных дифференциальных уравнений

$$\begin{aligned} \dot{x}(t) &= [\varepsilon_1 - \gamma_1 y(t)]x(t), \\ \dot{y}(t) &= [\gamma_2 x(t) - \varepsilon_2]y(t), \end{aligned} \quad (2.8)$$

хорошо изучена. Ее фазовый портрет представляет совокупность циклов вокруг неподвижной точки. Различные обобщения модели на случай запаздываний приведены в многочисленных исследованиях, см. [20, 48]. Так Вангерски и Каннингем рассмотрели следующую систему

$$\begin{aligned} \dot{x}(t) &= ax - cxy - bx^2, \\ \dot{y}(t) &= -ey + c_1 x(t - \tau)y(t - \tau). \end{aligned} \quad (2.9)$$

Запаздывание τ имеет смысл осредненного интервала времени между моментом гибели одной особи и моментом соответствующего увеличения числа взрослых хищников. У этой системы возможны различные модификации, когда запаздывание входит в одно или в оба уравнения.

Разнообразие математических моделей с запаздыванием в задаче хищник-жертва связано с тем, что само запаздывание может иметь самый разнообразный биологический смысл. Так в некоторых работах запаздывание входит в скорость вымирания хищников, вызванного голоданием (отрицательная реакция), и приводит к дестабилизации положения равновесия и возникновению устойчивого предельного цикла.

В некоторых моделях присутствует несколько запаздываний. Такова, например, модель Колесова

$$\dot{x}(t) = r_1(1 + a(1 - \frac{y}{K_2} - \frac{x(t - \tau_1)}{K_1})x,$$

$$\dot{y}(t) = r_2 \left(\frac{x}{K_1} - \frac{y(t - \tau_2)}{K_2} \right) y . \quad (2.10)$$

В этой модели сделано следующее основное допущение: считается, что сопротивление внешней среды есть внутренняя характеристика вида, т.е. оно не зависит от наличия или отсутствия хищника.

2.3. Модели эпидемий

Своеобразный класс задач порождает теория эпидемий [4], где вся популяция делится на группы восприимчивых индивидуумов, источников инфекции и изолированных индивидуумов. Так для системы

$$\begin{aligned} \dot{x}(t) &= bx(t - \tau)y(t - \tau) - cx , \\ \dot{y}(t) &= cx - bxy , \\ \dot{z}(t) &= b(xy - x(t - \tau)y(t - \tau)) , \end{aligned} \quad (2.11)$$

с условием $x + y + z = 1$ и интегральным ограничением

$$x + y = -a - b \int_{t-\tau}^t xy ds ,$$

выражающим наличием инкубационного периода, показано, что решение стремится к постоянной при $t \rightarrow \infty$.

Упрощенные модели могут выражаться интегральными уравнениями вида

$$x = \int_{t-\tau}^t f(s, x(s)) ds ,$$

при изучении которых запаздывание служит параметром. Эти уравнения эквивалентны уравнениям с сосредоточенным запаздыванием.

Отметим также более ранние модели Кука.

1-я модель Кука эпидемии [48])

$$\dot{x}(t) = b x(t - \tau) (1 - x(t)) - c x(t). \quad (2.12)$$

2-я модель Кука эпидемии [36])

$$\begin{cases} \dot{x}_1(t) = \lambda (x_2(t) - x_1(t)) \frac{x_1(t)}{x_2(t)} - (d + \varepsilon + \gamma) x_1(t), \\ \dot{x}_2(t) = b e^{-a x_2(t-\tau)} x_2(t - \tau) e^{-d_1 \tau} - d x_2(t) - \varepsilon x_1(t). \end{cases} \quad (2.13)$$

Два запаздывания имеет следующая модель

Модель эпидемии чумы [22])

$$\begin{cases} \dot{x}_1(t) = -x_1(t) x_2(t - \tau_1) + x_2(t - \tau_2), \\ \dot{x}_2(t) = x_1(t) x_2(t - \tau_1) - x_2(t), \\ \dot{x}_3(t) = x_2(t) - x_2(t - \tau_2). \end{cases} \quad (2.14)$$

2.4. Модели в иммунологии

Дифференциальные уравнения с запаздывающим аргументом широко используются при построении математических моделей в иммунологии. Так серии работ Б.Ф.Диброва, М.А.Лифшица и М.В.Волькенштейна [8] построены различные модели иммунной реакции, вызванной внедрением чужеродного материала – антигена (Аг). Скорость выработки специфических антител (Ат) можно считать пропорциональной наличному количеству плазматических клеток, определяемому числом В-лимфоцитов, стимулированных Аг на время τ_r ранее. Аналогично возможное образование клеток иммунной памяти определяется числом актов стимуляции В-лимфоцитов с запаздыванием τ_m . В результате одна из основных моделей представляет собой систему уравнений

$$\begin{aligned}\dot{x}(t) &= I - h^{-1}x - P x g + A_m x(t - \tau_m) g(t - \tau_m) \theta(t - \tau_m) , \\ \dot{g}(t) &= K g - Q g a , \\ \dot{a}(t) &= A_r x(t - \tau_r) g(t - \tau_r) \theta(t - \tau_r) - R g a - E a ,\end{aligned}\tag{2.15}$$

где x, g, a – соответственно количества В-лимфоцитов (со специфическими рецепторами к данному Аг), антигена и специфических Ат; $\theta(t)$ – функция Хевисайда; $P, I, h, A_m, A_r, K, Q, R, E$ – постоянные, характеризующие производство, гибель и взаимодействие x, g, a . Для этой модели и некоторых ее обобщений (в том числе и с распределенным запаздыванием) исследована зависимость динамики иммунного процесса от начальных данных и значений параметров, характеризующих иммунную систему и Аг. Показано, что одним из важнейших параметров, определяющих режим реакции, является продолжительность запаздывания τ_r в производстве Ат. При малом τ_r возможно стационарное сосуществование Аг и специфических Ат, увеличение же τ_r приводит к потере устойчивости стационарного состояния и к возникновению автоколебаний. Дальнейшее увеличение τ_r ведет к неограниченному размножению Аг, т.е. к гибели организма. Изменение продолжительности запаздывания может служить эффективным методом управления иммунными процессами. При малом τ_r велика вероятность инфекционного заболевания, тогда как незначительное увеличение τ_r может привести к существенному уменьшению этой вероятности. Чем больше τ_r , тем резче различия в ходе процесса в зависимости от дозы Аг.

2.5. Модель Марчука

Под термином “инфекционная болезнь” понимается отображение взаимоотношений, установившихся между членами биоценоза, один из которых способен благодаря механизмам патогенности существовать в другом, а другой благодаря механизмам защиты оказывать противодействие патогенному действию. Одним из основных способов защиты организма от инфекции является система иммунитета.

Суть иммунной реакции на вторжение генетически чужеродного материала (антигена), в том числе и возбудителя болезни, состоит в производстве специфических объектов (антител, клеток-киллеров), способных нейтрализовать или уничтожить антигены. Исходя из этого, будем рассматривать инфекционное заболевание как конфликт между популяцией возбудителей болезни и иммунной системы организма и в связи с этим выделим основные характеристики болезни.

В рамках модели процесс заболевания описывается следующим образом. В момент $t = t_0$ в организм проникает начальная популяция вирусов V_0 . Вирусы начинают размножаться в клетках органа-мишени и тем самым поражают его. Часть вирусов поступает в кровь, где взаимодействует с рецепторами иммунокомпетентных клеток (согласно модели — с антителами), в результате происходит стимуляция иммунной системы. Спустя время τ после стимуляции в организме появляются клоны плазматических клеток, вырабатывающие антитела, специфичные к вирусам. Антитела связывают вирусы, и от относительной кинетики этих процессов зависит исход болезни. Если вирусы успевают значительно поразить ткань органа, то ухудшается общее состояние организма и, как следствие, работы иммунной системы. Производство антител падает и снижается вероятность благоприятного исхода.

В модели рассматриваются следующие переменные: $V(t)$ — концентрация вирусов (размножающиеся патогенные антигены); $F(t)$ — концентрация антител (субстраты иммунной системы, нейтрализующие вирусы: иммуноглобулины, рецепторы клеток); $C(t)$ — концентрация плазматических клеток — популяция носителей и продуцентов антител (иммунокомпетентные клетки и иммуноглобулинопродуценты); $m(t)$ — относительная характеристика пораженного органа (масса, число пораженных клеток).

Модель сформулирована в виде следующей системы нелинейных дифференциальных уравнений [15]:

$$\begin{cases} \dot{V}(t) = (\beta - \gamma F(t))V(t), \\ \dot{C}(t) = \xi(m)\alpha V(t - \tau)F(t - \tau) - \mu_c(C(t) - C^*), \\ \dot{F}(t) = \rho C(t) - (\mu_f + \eta\gamma V(t))F(t), \\ \dot{m}(t) = \sigma V(t) - \mu_m m(t). \end{cases} \quad (2.16)$$

Система построена на соотношении баланса между “рождением” и “гибелью” каждого компонента процесса на малом интервале времени.

Первое уравнение описывает изменение числа вирусов в организме.

Второе уравнение описывает динамику плазматических клеток. Для построения уравнения используется простейшая гипотеза формирования каскадных популяций плазматических клеток.

Третье уравнение строится на основе баланса процессов производства антител, их взаимодействия с антигенами и естественной гибели.

Четвертое уравнение — это относительная характеристика поражения органа-мишени.

Система дополнена начальными данными:

$$\begin{aligned} V(t_0) &= V_0, C(t_0) = C_0, F(t_0) = F_0, m(t_0) = m_0; \\ F(t_0 + s) &= \hat{\varphi}(s), V(t_0 + s) = \hat{\psi}(s) \text{ при } -\tau \leq s < 0. \end{aligned} \quad (2.17)$$

Задача (2.16) – (2.17) названа простейшей математической моделью вирусного заболевания.

Простейшая математическая модель в данной трактовке допускает различные вариации, которые дают возможное объяснение некоторых важных особенностей функционирования иммунной системы, таких как формирование субклинических, острых и хронических процессов заболевания, выяснить роль температурных эффектов, изучить условия возникновения осложненных форм заболевания, а также механизм биостимуляции и другое.

2.6. LCMV-модель

Экспериментальная инфекция вирусами лимфоцитарного хориоменингита (LCMV) у мышей является фундаментальной экспериментальной системой в иммунологии. На её основе были открыты принципы презентирования вирусных антигенов и исследованы биологические механизмы взаимоотношений вирусной инфекции и иммунной системы хозяина. В частности, было установлено, что локализация, доза и кинетика вирусной инфекции являются ключевыми параметрами влияющими нелинейным образом на показатели иммунного ответа. Высокий темп репликации вирусов может приводить, в случае не- или слабо-цитопатических вирусов, к исчезновению вирусспецифических цитотоксических Т-лимфоцитов и хронизации инфекционного процесса. В работе [34] была построена математическая модель LCMV инфекции. В модели, описываемой системой нелинейных дифференциальных уравнений с запаздыванием, учтены процессы активации деления и дифференцировки цитотоксических Т-лимфоцитов (CTL), впервые рассмотрено состояние анергии и вирус-индуцированного апоптоза Т-лимфоцитов.

Уравнения модели описывают скорость изменения плотности популяции вирусов $V(t)$, CTL-предшественников — $E_p(t)$, CTL-эффекторов — $E(t)$ и уровень вирусной нагрузки $W(t)$:

$$\begin{cases} \dot{V}(t) = \beta V(t) (1 - V(t)/V_{max}) - \gamma_{VE} E(t) V(t), \\ \dot{E}_p(t) = \alpha_{E_p} (E_p^* - E_p(t)) + b_p g_p(W) V(t - \tau) E_p(t - \tau) - \\ \quad - \alpha_{AP} V(t - \tau_A) V(t) E_p(t), \\ \dot{E}(t) = b_d g_E(W) V(t - \tau) E_p(t - \tau) - b_{EV} V(t) E(t) - \\ \quad - \alpha_{AE} V(t - \tau_A) V(t) E(t) - \alpha_E E(t), \\ \dot{W}(t) = b_W V(t) - \alpha_W W(t), \end{cases}$$

где

$$g_p(W) = (1 + W/\theta_p)^{-2}; \quad g_E(W) = (1 + W/\theta_E)^{-2};$$

β — константа скорости репликации вирусов;

V_{max} — максимальная вирусная концентрация в селезенке;

γ_{VE} — константа нейтрализации вирусов за счет действия эффекторов;

α_{E_p} — константа скорости естественной гибели для CTL-предшественников;

E_p^* — гомеостатическая концентрация CTL-предшественников в селезенке до заражения;

b_p — константа скорости стимуляции CTL-предшественников;

θ_p — уровень вирусной нагрузки вызывающий снижение скорости деления CTL-предшественников в 2 раза;

τ — продолжительность цикла деления CTL-предшественников;

α_{AP} — константа скорости апоптоза для CTL-предшественников;

τ_A — продолжительность процесса формирования вирус-индуцированного апоптоза у CTL-предшественников;

b_d — константа скорости дифференциации CTL-предшественников;

- θ_E — уровень вирусной нагрузки вызывающий снижение в 2 раза скорости дифференцировки CTL-предшественников в эффекторы;
- b_{EV} — константа скорости гибели CTL-эффекторов вследствие взаимодействий с инфицированными клетками;
- α_{AE} — константа скорости апоптоза для CTL-эффекторов;
- α_E — константа скорости естественной гибели для CTL-эффекторов;
- b_W — константа скорости увеличения вирусной нагрузки;
- α_W — константа скорости уменьшения подавляющего эффекта вирусной нагрузки.

Первое уравнение описывает скорость изменения популяции вирусов в конкретном органе или во всём организме, которая зависит от скорости процессов размножения вирусов и их нейтрализации со стороны CTL-эффекторов.

Второе уравнение описывает динамику CTL предшественников. Учитывается вклад процессов поддержания гомеостаза предшественников, их активации и клонального деления, нелинейно зависящих от концентрации вирусов и уровня вирусной нагрузки, а также гибели связанной с вирус-индуцированным апоптозом.

Третье уравнение описывает динамику CTL-эффекторов. Рассматривается баланс процессов дифференцировки предшественников в эффекторы, нелинейно зависящий от концентрации вирусов и вирусной нагрузки и гибели связанной с вирус-индуцированным апоптозом. Явным образом учитывается запаздывание при описании процессов деления, дифференцировки и гибели CTL.

Четвёртое уравнение описывает изменение суммарной вирусной антигенной нагрузки в организме, которая влияет на скорости деления и дифференцировки CTL.

Параметры математической модели были идентифицированы по данным инфекции штаммом Dcsile ВЛХМ генетически идентичных мышей. Модель была исследована, используя численную технологию бифуркационного анализа для функционально-дифференциальных уравнений, в работе [34]. Были определены области в пространстве параметров модели, которым соответствуют различные динамические режимы сосуществования популяции вирусов и специфических CTL, например, устойчивое равновесие или периодические решения. Результаты исследования позволили теоретически предсказать условия персистирования вирусов ниже порога обнаружения (100 единиц на селезёнку) и изучить их вклад в поддержание иммунологической памяти, так называемого состояния «инфекционного иммунитета».

2.7. Некоторые другие модели

F-S-W модель (модель сосуществования двух микроорганизмов в питательной среде

$$\begin{cases} \dot{x}_0(t) = 1 - x_0(t) - x_1(t)f_1(x_0(t)) - x_2(t)f_2(x_0(t)), \\ \dot{x}_1(t) = [f_1(x_0(t - \tau_1))]x_1(t), \\ \dot{x}_2(t) = [f_2(x_0(t - \tau_2))]x_2(t). \end{cases} \quad (2.18)$$

Модель Велдона (модель хронической гранулоцитарной лейкемии [48])

$$\begin{cases} \dot{x}_1(t) = \frac{\alpha}{1 + \beta x_1(t - \tau)} - \frac{\lambda x_1(t)}{1 + \mu x_2^\delta(t)}, \\ \dot{x}_2(t) = \frac{\lambda x_1(t)}{1 + \mu x_2^\delta(t)} - \omega x_2(t). \end{cases} \quad (2.19)$$

Модель Оттесена (модель сердечно-сосудистой системы [49])

$$\begin{cases} \dot{x}_1(t) = -\frac{x_1(t)}{c_a R} + \frac{x_2(t)}{c_a R} + \frac{V_{str} x_3(t)}{c_a}, \\ \dot{x}_2(t) = \frac{x_1(t)}{c_v R} - \left(\frac{1}{c_v R} + \frac{1}{c_v r} \right) x_2(t), \\ \dot{x}_3(t) = \frac{\alpha_H (1 + (x_1(t - \tau)/\alpha_s)^{\beta_s})^{-1}}{1 + \gamma_H (1 + (\alpha_p/x_1(t))^{\beta_p})^{-1}} - \beta_H (1 + (\alpha_p/x_1(t))^{\beta_p})^{-1}. \end{cases} \quad (2.20)$$

Глава 3.

Одношаговые численные методы для ФДУ

Так как возникающие в биологии и медицине математические модели, описываемые ФДУ крайне редко поддаются аналитическому исследованию, основным инструментом становятся численные методы. Большую часть курса составляет изложение разработанного автором подхода к численному решению ФДУ.

3.1. Основные предположения

Рассмотрим систему функционально-дифференциальных уравнений

$$\dot{x}(t) = f(t, x(t), x_t(\cdot)), \quad (3.1)$$

($x_t(\cdot) = \{x(t+s), -\tau \leq s < 0\}$) с начальными условиями

$$x(t_0) = x_0, \quad (3.2)$$

$$x_{t_0}(\cdot) = \{y^0(s), -\tau \leq s < 0\}. \quad (3.3)$$

Здесь $f : [t_0, t_0 + \theta] \times \mathbb{R}^n \times Q[-\tau, 0] \rightarrow \mathbb{R}^n$; $\theta > 0$ - величина временного интервала.

При фиксированном t величину $x(t)$ будем называть *конечномерной*, а функцию-предысторию $x_t(\cdot)$ *бесконечномерной составляющей* фазового вектора. Особенность этой системы ФДУ состоит в том, что конечномерная составляющая отделена от бесконечномерной, в отличие от обычно рассматриваемых ФДУ [24].

В дальнейшем будем предполагать выполненными следующие условия:

Условие 3.1. Отображение f в своей области определения *непрерывно по сдвигу* в смысле определения [13].

Условие 3.2. Отображение f в своей области определения *липшицево* по x и y , т.е. найдутся постоянные L и M такие, что для всех $t \in [t_0, t_0 + \theta]$; $x^{(1)}, x^{(2)} \in \mathbb{R}^n$; $y^{(1)}(\cdot), y^{(2)}(\cdot) \in Q[-\tau, 0]$ выполняется

$$\begin{aligned} & \|f(t, x^{(1)}, y^{(1)}(\cdot)) - f(t, x^{(2)}, y^{(2)}(\cdot))\| \leq \\ & \leq L\|x^{(1)} - x^{(2)}\| + M\|y^{(1)}(\cdot) - y^{(2)}(\cdot)\|_Q. \end{aligned}$$

Из теоремы существования и единственности см, например, [13, 17] следует, что найдется $\hat{\Delta} > 0$ такое, что решение задачи (3.1) - (3.3) существует и единственно на $[t_0, t_0 + \hat{\Delta}]$.

Условие 3.2 можно ослабить, требуя лишь локальную липшицевость.

Ниже приведены некоторые численные модели решения систем ФДУ. Проиллюстрируем особенности этих методов на простейшей модели – методе Эйлера.

3.2. Численный метод Эйлера с кусочно-постоянной интерполяцией

Зададим на $[t_0, t_0 + \theta]$ временную сетку $t_l = t_0 + l\Delta$, $l = 0, 1, \dots, N$ с равномерным шагом $\Delta = \theta/N$, где N — целое число. Для простоты будем считать, что $\tau/\Delta = m$ — целое число. Введем *дискретную численную модель* системы (3.1), обозначив приближение точного решения $x(t_l) = x_l$ в точке t_l через $u_l \in \mathbb{R}^n$.

В отличие от обыкновенных дифференциальных уравнений, в системе (3.1) функционал правой части f определён на функциях-предысториях, поэтому задание дискретной предыстории недостаточно для построения адекватной системе (3.1) численной модели. Для того чтобы определить функционал f на приближенном решении, необходима интерполяция. Простейший способ — кусочно-постоянная интерполяция:

$$u(t) = \begin{cases} u_i, & t \in [t_i, t_{i+1}), \quad i = 0, 1, \dots, l, \\ y^0(t - t_0), & t \in [t_0 - \tau, t_0). \end{cases} \quad (3.4)$$

Отметим, что в момент t_l определена функция-предыстория $u_{t_l}(\cdot) \in Q[-\tau, 0)$, если задана дискретная предыстория модели u_i , $i \leq l$.

Методом Эйлера с кусочно-постоянной интерполяцией дискретной предыстории модели назовем пошаговую модель

$$u_0 = x_0; \quad (3.5)$$

$$u_{l+1} = u_l + \Delta f(t_l, u_l, u_{t_l}(\cdot)), \quad l = 0, \dots, N-1, \quad (3.6)$$

где $u_{t_l}(\cdot) = \{u(t_l + s), -\tau \leq s < 0\}$ — предыстория модели, определенная интерполяцией (3.4).

Будем говорить, что метод *сходится*, если

$$\max_{1 \leq l \leq N} \|u_l - x(t_l)\| \rightarrow 0 \text{ при } N \rightarrow \infty,$$

и имеет *порядок сходимости* p , если найдется постоянная C такая, что $\|u_l - x(t_l)\| \leq C\Delta^p$ для всех $l = 1, \dots, N$ и константа C не зависит от N .

Цель дальнейших утверждений — показать сходимость метода Эйлера (3.4) — (3.6) и доказать, что он имеет первый порядок сходимости.

Лемма 3.1. *Если точное решение $x(t)$ — непрерывно дифференцируемая функция, то результат кусочно-постоянной интерполяции обладает свойством*

$$\|x_{t_l}(\cdot) - u_{t_l}(\cdot)\|_Q \leq \max_{l-m \leq i \leq l} \|x_i - u_i\| + C_1 \Delta, \quad l = 0, \dots, N. \quad (3.7)$$

Доказательство. Пусть $\hat{x}(t)$ — кусочно-постоянная функция, определяемая соотношением

$$\hat{x}(t) = (x(t_i), \quad t \in [t_i, t_{i+1})).$$

Тогда

$$\begin{aligned} \|x_{t_l}(\cdot) - u_{t_l}(\cdot)\|_Q &= \sup_{t_l - \tau \leq t \leq t_l} \|x(t) - u(t)\| \leq \\ &\leq \sup_{t_l - \tau \leq t \leq t_l} \|x(t) - \hat{x}(t)\| + \sup_{t_l - \tau \leq t \leq t_l} \|\hat{x}(t) - u(t)\| \leq \end{aligned}$$

$$\leq \max_{l-m \leq i \leq l} \|x_i - u_i\| + C_1 \Delta,$$

где C_1 – максимум модуля производной функции $x(t)$ на отрезке $[t_0, t_0 + \theta]$. \square

Невязкой $\psi(t_l)$ назовем функцию

$$\psi(t_l) = \frac{x_{l+1} - x_l}{\Delta} - f(t_l, x_l, x_{t_l}(\cdot)).$$

Лемма 3.2. *Если $x(t)$ дважды непрерывно дифференцируемая функция, то невязка метода Эйлера имеет первый порядок малости по Δ .*

Доказательство. Разложим $x(t_{l+1})$ по формуле Тейлора в окрестности точки t_l :

$$x(t_{l+1}) = x(t_l) + \Delta f(t_l, x_l, x_{t_l}(\cdot)) + \ddot{x}(c) \frac{\Delta^2}{2}, \quad c \in [t_l, t_{l+1}],$$

тогда для всякого $l = 0, \dots, N-1$ выполняется

$$\|\psi(t_l)\| \leq C_2 \Delta,$$

где C_2 – половина максимума модуля второй производной функции $x(t)$ на отрезке $[t_0, t_0 + \theta]$. \square

Теорема 3.1. *Если выполняются условия леммы 3.2, а также условия 3.1 и 3.2, то метод Эйлера с кусочно-постоянной интерполяцией сходится, причем порядок сходимости – первый.*

Доказательство. Обозначим $\epsilon_l = \|x_l - u_l\|$ и выразим величину ϵ_{l+1} через ϵ_l , при этом x_{l+1} подставим из определения невязки, а u_{l+1} из пошаговой формулы (3.6):

$$\begin{aligned} \epsilon_{l+1} &= \|x_{l+1} - u_{l+1}\| = \\ &= \|x_l + \Delta f(t_l, x_l, x_{t_l}(\cdot)) + \Delta \psi(t_l) - u_l - \Delta f(t_l, u_l, u_{t_l}(\cdot))\| \leq \\ &\leq \epsilon_l + \Delta \|f(t_l, x_l, x_{t_l}(\cdot)) - f(t_l, u_l, u_{t_l}(\cdot))\| + \Delta \|\psi(t_l)\|. \end{aligned}$$

Пользуясь условием 3.2 (липшицевостью функционала f) и утверждением леммы 3.2, получаем оценку

$$\epsilon_{l+1} \leq \epsilon_l + \Delta(L\|x_l - u_l\| + M\|x_{t_l}(\cdot) - u_{t_l}(\cdot)\|_Q) + C_2 \Delta^2,$$

откуда, с использованием леммы 3.1, имеем

$$\epsilon_{l+1} \leq \epsilon_l(1 + L\Delta) + \Delta M \max_{l-m \leq i \leq l} \|\epsilon_i\| + \Delta^2(MC_1 + C_2). \quad (3.8)$$

Индукцией по l докажем оценку

$$\epsilon_l \leq (1 + \Delta(L + M + 1))^l (MC_1 + C_2) \Delta. \quad (3.9)$$

База индукции выполняется, так как $\epsilon_0 = 0$.

Шаг индукции. Пусть оценка (3.9) выполняется для индексов $\leq l$, покажем её справедливость для $l+1$.

Пусть тах в правой части оценки (3.8) достигается на индексе $i_0 \leq l$ тогда, применяя индуктивное предположение к ϵ_l и ϵ_{i_0} , получаем

$$\epsilon_{l+1} \leq \epsilon_l(1 + L\Delta) + \Delta M \epsilon_{i_0} + \Delta^2(MC_1 + C_2) \leq$$

$$\begin{aligned}
&\leq (1 + \Delta(L + M + 1))^l (MC_1 + C_2)\Delta(1 + L\Delta) + \\
&+ \Delta M(1 + \Delta(L + M + 1))^{i_0} (MC_1 + C_2)\Delta + \Delta^2(MC_1 + C_2) \leq \\
&\leq (1 + \Delta(L + M + 1))^l (MC_1 + C_2)\Delta(1 + L\Delta + M\Delta + \Delta).
\end{aligned}$$

Оценка (3.9) доказана .

Так как $N = \frac{\theta}{\Delta}$, то для всех $l = 0, \dots, N$ имеем оценку

$$\begin{aligned}
\epsilon_l &\leq (1 + \Delta(L + M + 1))^{\frac{\theta}{\Delta}} (MC_1 + C_2)\Delta \leq \\
&\leq e^{(L+M+1)\theta} (MC_1 + C_2)\Delta.
\end{aligned}$$

Эта оценка содержит утверждение теоремы. \square

Описанный метод Эйлера с кусочно-постоянной интерполяцией является простейшим из сходящихся методов. Дальнейшее улучшение свойств сходимости возможно, во-первых, за счет усложнения интерполяции, во-вторых, за счет усложнения пошаговой модели.

3.3. Способы интерполяции и экстраполяции предыстории дискретной модели

Пусть на отрезке $[t_0, t_0 + \theta]$ задана временная сетка, такая же, как и в предыдущем разделе. *Дискретной предысторией модели в момент t_l* назовем множество из $m + 1$ вектора:

$$\{u_i\}_l = \{u_i \in \mathbb{R}^n, \quad l - m \leq i \leq l\}.$$

Этот набор значений определяет в момент t_l дальнейшее поведение модели.

Определение 3.1. *Оператором интерполирования I дискретной предыстории модели назовем отображение $I : \{u_i\}_l \rightarrow u(\cdot) \in Q[t_l - \tau, t_l]$.*

Определение 3.2. *Оператор I имеет порядок погрешности p на точном решении, если существуют константы C_1, C_2 такие, что для всех $l = 0, 1, \dots, N$ и $t \in [t_l - \tau, t_l]$ выполняется неравенство*

$$\|x(t) - u(t)\| \leq C_1 \max_{l-m \leq i \leq l} \|u_i - x_i\| + C_2 \Delta^p. \quad (3.10)$$

Примером оператора интерполяции второго порядка на достаточно гладком точном решении служит интерполяция ломаными (кусочно-линейная интерполяция):

$$\begin{aligned}
I : \{u_i\}_l &\rightarrow u(t) = \\
&= \begin{cases} ((t - t_i)u_{i+1} + (t_{i+1} - t)u_i)\frac{1}{\Delta}, & t \in [t_i, t_{i+1}], \quad i = 0, \dots, l-1, \\ y^0(t_0 - t), & t \in [t_l - \tau, t_0]. \end{cases} \quad (3.11)
\end{aligned}$$

Рассмотрим более общий способ интерполяции функциями, кусочно составленными из многочленов p -й степени, где p — произвольное натуральное число. Такие функции называются вырожденными сплайнами. Без ограничения общности будем

предполагать, что $\frac{m}{p} = k$ — целое, в противном случае можно выбирать m кратным p .

Пусть $t_l \in [t_0, t_0 + \theta]$, $t_i = t_0 + i\Delta$, $i = l - m, \dots, l$ — узлы предыстории, u_i , $i = l - m, \dots, l$ — соответствующая дискретная предыстория модели. Заметим, что может выполняться соотношение $t_i < t_0$, в этом случае $u_i = y^0(t_0 - t_i)$; в случае $t_i = t_0$: $u_0 = x_0$.

Разобьем отрезок $[t_l - \tau, t_l]$ справа на k подотрезков $[t_{l_{i+1}}, t_{l_i}]$, $i = 0, 1, \dots, k - 1$, длиной $p\Delta$ таким образом, что $t_{l_0} = t_l$, $t_{l_1} = t_{l-p}$, \dots . На каждом отрезке $[t_{l_{i+1}}, t_{l_i}]$ построим интерполяционный многочлен $L_p(t) = L_p^i(t)$ по данным $u_{l_i-p}, u_{l_i-p+1}, \dots, u_{l_i}$:

$$L_p^i(t) = \sum_{n=0}^p u_{l_i-n} \prod_{j=l_i-p; j \neq l_i-n}^{l_i} \frac{t - t_j}{t_{l_i-n} - t_j}.$$

Определение 3.3. Оператором интерполирования предыстории модели вырожденными сплайнами p -го порядка назовем

$$I : \{u_i\}_l \rightarrow u(t) = \begin{cases} L_p^i(t), & t_{l_{i+1}} \leq t < t_{l_i}, \quad i = 0, \dots, l - 1, \\ y^0(t_0 - t), & t \in [t_l - \tau, t_0]. \end{cases} \quad (3.12)$$

Теорема 3.2. Пусть точное решение $x(t)$ является $p+1$ раз непрерывно дифференцируемым на отрезке $[t_0 - \tau, t_0 + \theta]$, тогда оператор интерполяции вырожденными сплайнами p -го порядка имеет погрешность интерполяции $p+1$.

Доказательство. Возьмем произвольное $l = 0, \dots, N$ и произвольное $t \in [t_l - \tau, t_l]$. Пусть t лежит в подотрезке $[t_{l_{i+1}}, t_{l_i}]$. Тогда

$$\|x(t) - u(t)\| = \|x(t) - L_p^i(t)\| \leq \|x(t) - \hat{L}_p^i(t)\| + \|L_p^i(t) - \hat{L}_p^i(t)\|,$$

где $\hat{L}_p^i(t)$ — интерполяционный многочлен p -й степени, построенный по значениям x_j на отрезке $[t_{l_{i+1}}, t_{l_i}]$:

$$\hat{L}_p^i(t) = \sum_{n=0}^p x_{l_i-n} \prod_{j=l_{i+1}; j \neq l_i-n}^{l_i} \frac{t - t_j}{t_{l_i-n} - t_j}.$$

Погрешность интерполяции точного решения вычисляется по формуле [19, с. 133]

$$\|R_p(t)\| = \|x(t) - \hat{L}_p^i(t)\| = \left\| \frac{x^{(p+1)}(\xi)}{(p+1)!} \prod_{j=0}^p (t - t_{l_i-j}) \right\|,$$

где $\xi \in [t_{l_{i+1}}, t_{l_i}]$. Таким образом, если $t \in [t_{l_{i+1}}, t_{l_i}]$, то

$$\left\| \prod_{j=0}^p (t - t_{l_i-j}) \right\| \leq \Delta \cdot \Delta \cdot (2\Delta) \cdot \dots \cdot (p\Delta) = \Delta^{p+1} p!$$

и

$$\|R_p(t)\| \leq \frac{M_{p+1}}{(p+1)!} p! \Delta^{p+1} = \frac{M_{p+1}}{p+1} \Delta^{p+1}, \quad (3.13)$$

где $M_{p+1} = \max_{t_0-\tau \leq t \leq t_0+\theta} \|x^{(p+1)}(t)\|$. Кроме того, имеем оценку

$$\begin{aligned} \|L_p^i(t) - \hat{L}_p^i(t)\| &= \sum_{n=0}^p \|u_{l_i-n} - x_{l_i-n}\| \prod_{j=l_i-p; j \neq l_i-n}^{l_i} \frac{|t - t_j|}{|t_{l_i-n} - t_j|} \leq \\ &\leq \max_{l_i-p \leq n \leq l_i} \|u_n - x_n\| (p+1)!. \end{aligned} \quad (3.14)$$

Из (3.13) и (3.14) получаем

$$\|x(t) - L_p^i(t)\| \leq C_1 \max_{l_i-k \leq n \leq l_i} \|u_n - x_n\| + C_2 \Delta^{p+1}. \quad (3.15)$$

Здесь $C_1 = (p+1)!$, $C_2 = \frac{M_{p+1}}{p+1}$. □

Рассмотрим модификацию интерполяции вырожденными сплайнами, которую назовем *интерполяцией по ближайшим узлам*.

Пусть t_l фиксировано, $1 \leq p \leq m$ — натуральное число. На каждом из отрезков $[t_i, t_{i+1}]$, $i = l-m, \dots, l-1$, построим многочлены $L_p^i(t)$ следующим образом.

Определим $p+1$ узел интерполяции t_{i_j} соотношениями $t_{i_0} = t_i$, $t_{i_1} = t_{i+1}$, $t_{i_2} = t_{i-1}$, $t_{i_3} = t_{i+2}$ и так далее, если $t_{i_j} \leq t_l$. Если этот алгоритм выбора узлов интерполяции приводит для какого-то j к соотношению $t_{i_j} > t_l$, т.е. к выходу узла за t_l вправо, то в качестве t_{i_j} берем ближайший неиспользованный узел слева от $[t_i, t_{i+1}]$. Например, если $t_{i_3} = t_{i+2} > t_l$ (это будет когда $i = l-1$), то в качестве t_{i_3} надо взять t_{i-2} .

Переобозначив соответствующим образом предысторию модели u_{i_j} , $j = 0, \dots, p$, построим на $[t_i, t_{i+1}]$ интерполяционный многочлен

$$L_p^i(t) = \sum_{j=0}^p u_{i_j} \prod_{k=0; k \neq j}^p \frac{t - t_k}{t_{i_j} - t_k}.$$

Если t_{i_j} выходит за пределы отрезка $[t_0, t_0 + \theta]$ слева, т.е. $t_{i_j} < t_0$, то $u_{i_j} = y^0(t_0 - t_{i-j})$; в случае $t_{i_j} = t_0$: $u_0 = x_0$.

Определим оператор интерполирования предыстории модели

$$I : \{u_i\}_l \rightarrow u(t) = \begin{cases} L_p^i(t), & t_i \leq t < t_{i+1}, \quad i = 0, \dots, l-1, \\ y^0(t_0 - t), & t \in [t_l - \tau, t_0]. \end{cases}$$

Заключение теоремы 3.2 остается справедливым и для этого способа интерполяции, более того, в некоторых ситуациях оценка в доказательстве этой теоремы может быть улучшена.

Лемма 3.3. Пусть выполняются условия теоремы 3.2, и пусть $[t_i, t_{i+1}] \subset [t_l - \tau, t_l]$, причем $t_i + (p-1)\Delta \leq t_l$ при нечетном p или $t_i + (p-2)\Delta \leq t_l$ при четном p . Тогда оператор интерполирования по ближайшим узлам обладает свойством: для любого $t \in [t_i, t_{i+1}]$

$$\|x(t) - u(t)\| \leq \hat{C}_1 \max_{l-m \leq n \leq l} \|u_n - x_n\| + \hat{C}_2 \Delta^{p+1},$$

где при $p = 2k + 1$

$$\hat{C}_1 = (2k+2)((k+1)!)^2, \quad \hat{C}_2 = \frac{((k+1)!)^2 M_{p+1}}{(2k+2)!},$$

при $p = 2k$

$$\hat{C}_1 = (2k+1)(k!(k+1)!), \quad \hat{C}_2 = \frac{k!(k+1)M_{p+1}}{(2k+1)!}.$$

Доказательство. Возьмем произвольное $t \in [t_i, t_{i+1}] \subset [t_l - \tau, t_l]$. Из условий леммы следует, что для всех узлов интерполяции выполняется $t_{i_j} \leq t_l$. Тогда

$$\|x(t) - u(t)\| = \|x(t) - L_p^i(t)\| \leq \|x(t) - \hat{L}_p^i(t)\| + \|L_p^i(t) - \hat{L}_p^i(t)\|,$$

где $\hat{L}_p^i(t)$ — интерполяционный многочлен p -й степени, построенный по значениям $x_{i_j} = x(t_{i_j})$:

$$L_p^i(t) = \sum_{j=0}^p x_{i_j} \prod_{n=0; n \neq j}^p \frac{t - t_n}{t_{i_j} - t_n}.$$

Погрешность интерполяции точного решения вычисляется по формуле

$$\|R_p(t)\| = \|x(t) - \hat{L}_p^i(t)\| = \left\| \frac{x^{(p+1)}(\xi)}{(p+1)!} \prod_{j=0}^p (t - t_{i_j}) \right\|,$$

где $\xi \in [t_{l-m}, t_l]$.

Рассмотрим два случая:

а) p нечетное, т.е. $p = 2k + 1$.

Если $t \in [t_i, t_{i+1}]$, то

$$\left\| \prod_{j=0}^p (t - t_{i_j}) \right\| \leq \Delta \cdot \Delta \cdot (2\Delta)(2\Delta) \cdots ((k+1)p\Delta)((k+1)p\Delta) = \Delta^{p+1}((k+1)!)^2$$

и

$$\|R_p(t)\| \leq \frac{M_{p+1}((k+1)!)^2}{2k+2} \Delta^{p+1},$$

где $M_{p+1} = \max_{t_0 - \tau \leq t \leq t_0 + \theta} \|x^{(p+1)}(t)\|$. Кроме того, имеем оценку

$$\begin{aligned} \|L_p^i(t) - \hat{L}_p^i(t)\| &= \sum_{j=0}^p \|u_{i_j} - x_{i_j}\| \prod_{n=0; n \neq i_j}^p \frac{|t - t_n|}{|t_{i_j} - t_n|} \leq \\ &\leq \max_{l-m \leq n \leq l} \|u_n - x_n\| (p+1)((k+1)!)^2. \end{aligned}$$

Отсюда получаем утверждение леммы для нечетного p .

б) p четное, т.е. $p = 2k$. В этом случае изменятся оценки

$$\left\| \prod_{j=0}^p (t - t_{i_j}) \right\| \leq \Delta \cdot \Delta \cdot (2\Delta)(2\Delta) \cdots (kp\Delta)(kp\Delta)((k+1)p\Delta) = \Delta^{p+1}k!(k+1)!$$

и

$$\prod_{n=0; n \neq i_j}^p \frac{|t - t_n|}{|t_{i_j} - t_n|} \leq k!((k+1)!).$$

Отсюда получаем утверждение леммы для четного p . \square

Соответствующие оценки можно получить и в ситуации, когда лишь часть узлов интерполяции удовлетворяет условиям $t_{i_j} \leq t_l$, эти оценки также дают меньшие константы, чем C_1 и C_2 в общей ситуации, определенной в теореме 3.2.

Кроме интерполяции вырожденными сплайнами в рассматриваемых ниже методах можно применять интерполяцию многочленами Эрмита с кратными узлами, используя условие $\dot{u}_l = f(t_l, u_l, u_{t_l})$, а также другие способы интерполяции дискретной предыстории модели.

Во многих рассматриваемых методах (типа Рунге–Кутты, неявных многошаговых и др.) необходимо в момент t_l знать предысторию модели $u_{t_l+a\Delta}(\cdot)$ при $a > 0$, т.е. произвести экстраполяцию модели на отрезок $[t_l, t_l + a\Delta]$.

Определение 3.4. Для любого $a > 0$ оператором экстраполирования E предыстории модели назовем отображение $E : \{u_i\}_l \rightarrow u(\cdot) \in Q[t_l, t_l + a\Delta]$.

Определение 3.5. Будем говорить, что экстраполяция предыстории модели имеет порядок погрешности p на точном решении, если существуют константы C_3, C_4 такие, что для всех $a > 0$ и всех $l = 0, 1, \dots, N-1$ и $t \in [t_l, t_l + a\Delta]$ выполняется неравенство

$$\|x(t) - u(t)\| \leq C_3 \max_{l-m \leq i \leq l} \|u_i - x_i\| + C_4(\Delta)^p. \quad (3.16)$$

Один из способов задания оператора экстраполяции – экстраполяция продолжением интерполяционного многочлена

$$E : \{u_i\}_l \rightarrow u(t) = L_p^0(t), \quad t \in [t_l, t_l + a\Delta], \quad (3.17)$$

где $L_p^0(t)$ – интерполяционный многочлен p -ой степени, построенный по значениям u_j на отрезке $[t_{l-p}, t_l]$:

$$L_p^0(t) = \sum_{n=0}^p u_{l-n} \prod_{j=l-p; j \neq l-n}^l \frac{t - t_j}{t_{l-n} - t_j}.$$

Теорема 3.3. Пусть точное решение $x(t)$ является $p+1$ раз непрерывно дифференцируемым на отрезке $[t_0 - \tau, t_0 + \theta]$, тогда оператор экстраполяции продолжением интерполяционного многочлена p -й степени имеет погрешность экстраполяции порядка $p+1$.

Доказательство. Возьмем произвольное $t \in [t_l, t_l + a\Delta]$. Тогда

$$\|x(t) - u(t)\| = \|x(t) - L_p^0(t)\| \leq \|x(t) - \hat{L}_p^0(t)\| + \|L_p^0(t) - \hat{L}_p^0(t)\|,$$

где $\hat{L}_p^0(t)$ – интерполяционный многочлен p -й степени, построенный по значениям x_j на отрезке $[t_{l-p}, t_l]$.

Тогда

$$\|x(t) - \hat{L}_p^0(t)\| \leq \frac{M_{p+1}}{(p+1)!} \prod_{j=0}^p (t - t_{l-j}) \leq$$

$$\leq \frac{M_{p+1}}{(p+1)!} \prod_{j=0}^p (a\Delta + j\Delta) \leq C_4(\Delta)^{p+1},$$

где

$$C_4 = \frac{M_{p+1}}{(p+1)!} \prod_{j=0}^p (a+j).$$

Кроме того,

$$\begin{aligned} \|L_p^0(t) - \hat{L}_p^0(t)\| &\leq \max_{l-p \leq n \leq l} \|u_n - x_n\| \sum_{n=0}^p \prod_{j=0; j \neq n}^p \frac{(a\Delta + j\Delta)}{\Delta} \leq \\ &\leq C_3 \max_{l-p \leq n \leq l} \|u_n - x_n\|, \end{aligned}$$

где

$$C_3 = (p+1) \prod_{j=0}^p (a+j).$$

Теорема доказана. \square

Еще один способ экстраполяции предыстории модели основывается на экстраполяции с помощью непрерывных методов (см. ниже). При этом порядок погрешности экстраполяции p совпадает с локальным порядком аппроксимации непрерывного метода. Так, например, экстраполяция по методу Эйлера

$$E : \{u_i\}_l \rightarrow u(t) = u_l + (t - t_l)f(t_l, u_l, u_{t_l}), t \in [t_l, t_l + a\Delta]$$

имеет порядок погрешности $p = 2$ при условии, что $x(\cdot)$ достаточно гладкое.

Иногда имеет смысл объединить операторы интерполяции и экстраполяции.

Определение 3.6. Для любого $a > 0$ оператором интерполяции-экстраполяции IE предыстории модели назовем отображение $IE : \{u_i\}_l \rightarrow u(\cdot) \in Q[t_l - \tau, t_l + a\Delta]$.

Определение 3.7. Будем говорить, что оператор IE имеет порядок погрешности p на точном решении, если существуют константы C_5, C_6 такие, что для всех $a > 0, l = 0, 1, \dots, N-1$ и $t \in [t_l - \tau, t_l + a\Delta]$ выполняется неравенство

$$\|x(t) - u(t)\| \leq C_5 \max_{l-m \leq i \leq l} \|u_i - x_i\| + C_6(\Delta)^p. \quad (3.18)$$

3.4. Явные методы типа Рунге–Кутты

Предположим, что заданы операторы интерполяции I и экстраполяции E .

Определение 3.8. Для натурального k назовем k -этапным явным методом типа Рунге–Кутты – ЯРК (с интерполяцией I и экстраполяцией E) численную модель вида

$$u_0 = x_0, \quad (3.19)$$

$$u_{l+1} = u_l + \Delta \sum_{i=1}^k \sigma_i h_i(u_l, u_{t_l}(\cdot)), \quad l = 0, \dots, N-1, \quad (3.20)$$

$$h_1(u_l, u_{t_l}(\cdot)) = f(t_l, u_l, u_{t_l}(\cdot)), \quad (3.21)$$

$$h_i(u_l, u_{t_l}(\cdot)) = f(t_l + a_i \Delta, u_l + \Delta \sum_{j=1}^{i-1} b_{ij} h_j(u_l, u_{t_l}(\cdot)), u_{t_l + a_i \Delta}(\cdot)). \quad (3.22)$$

Здесь предыстория модели определяется соотношениями

$$u_t(s) = \begin{cases} y^0(t + s - t_0) & \text{при } t + s < t_0, \\ I(\{u_i\}_l) & \text{при } t_l - \tau \leq t + s < t_l, \\ E(\{u_i\}_l) & \text{при } t_l \leq t + s \leq t_l + a\Delta, \end{cases} \quad (3.23)$$

$$a = \max\{|a_i|, 1 \leq i \leq k\}.$$

Числа a_i, σ_i, b_{ij} называются *коэффициентами* метода. Будем обозначать $\sigma = \max\{|\sigma_i|\}, b = \max\{|b_{ij}|\}$.

Изучим порядок сходимости (в смысле определения раздела 3.2) ЯРК-метода.

Определение 3.9. *Невязкой (погрешностью аппроксимации) ЯРК-метода назовем функцию*

$$\psi(t_l) = \frac{x_{l+1} - x_l}{\Delta} - \sum_{i=1}^k \sigma_i h_i(x_l, x_{t_l}(\cdot)).$$

Заметим, что невязка определена на точном решении $x(t)$ и не зависит от интерполяции и экстраполяции.

Определение 3.10. *Будем говорить, что невязка имеет порядок p , если найдется постоянная C такая, что $\|\psi(t_l)\| \leq C\Delta^p$ для всех $l = 0, 1, \dots, N-1$.*

Теорема 3.4. *Если метод (3.19) - (3.23) имеет невязку порядка $p_1 > 0$, интерполяция предыстории модели имеет порядок $p_2 > 0$, экстраполяция предыстории модели имеет порядок $p_3 > 0$, то метод сходится, причем порядок сходимости ЯРК-метода p не меньше минимума из p_1, p_2, p_3 .*

Для доказательства этой теоремы сходимости рассмотрим два вспомогательных утверждения.

Лемма 3.4. *Функционалы h_i , определяемые в (3.21) - (3.22), липшицевы в следующем смысле: найдутся такие константы L_i и M_i , что*

$$\begin{aligned} & \|h_i(u_l, u_{t_l}(\cdot)) - h_i(x_l, x_{t_l}(\cdot))\| \leq \\ & \leq L_i \|u_l - x_l\| + M_i \sup_{t_l - \tau \leq t \leq t_l + a\Delta} \|u(t) - x(t)\|. \end{aligned}$$

Доказательство проведем индукцией по i .

При $i = 1$ функционал $h_1(u_l, u_{t_l}(\cdot)) = f(t_l, u_l, u_{t_l}(\cdot))$ и по условию 3.2 липшицев, причем $L_1 = L, M_1 = M$.

Предположим, что для индексов $j \leq i-1$ функционалы h_j липшицевы с константами L_i, M_i . Докажем липшицевость функционала h_i :

$$\|h_i(u_l, u_{t_l}(\cdot)) - h_i(x_l, x_{t_l}(\cdot))\| =$$

$$\begin{aligned}
&= \|f(t_l + a_i \Delta, u_l + \Delta \sum_{j=1}^{i-1} b_{ij} h_j(u_l, u_{t_l}(\cdot)), u_{t_l + a_i \Delta}(\cdot)) - \\
&\quad - f(t_l + a_i \Delta, x_l + \Delta \sum_{j=1}^{i-1} b_{ij} h_j(x_l, x_{t_l}(\cdot)), x_{t_l + a_i \Delta}(\cdot))\| \leq \\
&\leq L \|u_l - x_l\| + L \Delta b \sum_{j=1}^{i-1} \|h_j(u_l, u_{t_l}(\cdot)) - h_j(x_l, x_{t_l}(\cdot))\| + \\
&\quad + M \sup_{t_l - \tau \leq t \leq t_l + a \Delta} \|u(t) - x(t)\| \leq \\
&\leq L \|u_l - x_l\| + L \Delta b \sum_{j=1}^{i-1} (L_j \|u_l - x_l\| + \\
&\quad + M_j \sup_{t_l - \tau \leq t \leq t_l + a \Delta} \|u(t) - x(t)\|) + M \sup_{t_l - \tau \leq t \leq t_l + a \Delta} \|u(t) - x(t)\|.
\end{aligned}$$

Таким образом, функционал h_i липшицев с константами

$$L_i = L + L \Delta b \sum_{j=1}^{i-1} L_j, \quad M_i = M + L \Delta b \sum_{j=1}^{i-1} M_j.$$

□

Лемма 3.5. Если оператор интерполяции обладает порядком p_2 , оператор экстраполяции обладает порядком p_3 , то найдутся такие константы \hat{C}_1 и \hat{C}_2 , что выполняется неравенство

$$\begin{aligned}
&\|h_i(u_l, u_{t_l}(\cdot)) - h_i(x_l, x_{t_l}(\cdot))\| \leq \\
&\leq \hat{C}_1 \max_{l-m \leq i \leq l} \|u_i - x_i\| + \hat{C}_2 \Delta^{\min\{p_2, p_3\}}.
\end{aligned}$$

Доказательство. Заметим, что из того, что оператор интерполяции I обладает порядком p_2 , а оператор экстраполяции E обладает порядком p_3 , следует, что оператор IE имеет порядок $p = \min\{p_2, p_3\}$. Тогда из (3.18) и предыдущей леммы вытекает

$$\begin{aligned}
&\|h_i(u_l, u_{t_l}(\cdot)) - h_i(x_l, x_{t_l}(\cdot))\| \leq \\
&\leq L_i \|u_l - x_l\| + M_i \sup_{t_l - \tau \leq t \leq t_l + a \Delta} \|u(t) - x(t)\| \leq \\
&\leq L_i \|u_l - x_l\| + M_i (C_5 \max_{l-m \leq i \leq l} \|u_i - x_i\| + \hat{C}_6 \Delta^{\min\{p_2, p_3\}}).
\end{aligned}$$

Взяв

$$\hat{C}_1 = \max_{1 \leq i \leq k} M_i C_5 + \max_{1 \leq i \leq k} L_i, \quad \hat{C}_2 = \max_{1 \leq i \leq k} M_i C_6,$$

получим утверждение леммы.

□

Доказательство теоремы 3.4. Обозначим $\epsilon_l = \|x_l - u_l\|$ и выразим величину ϵ_{l+1} через ϵ_l , при этом x_{l+1} подставим из определения невязки, а u_{l+1} – из пошаговой формулы (3.20). С использованием леммы 3.5 получаем:

$$\begin{aligned}\epsilon_{l+1} &= \|x_{l+1} - u_{l+1}\| = \\ &= \|x_l + \Delta\psi(t_l) - u_l + \Delta \sum_{i=1}^k \sigma_i (h_i(x_l, x_{t_l}(\cdot)) - h_i(u_l, u_{t_l}(\cdot)))\| \leq \\ &\leq \epsilon_l + \Delta\|\psi(t_l)\| + \Delta \sum_{i=1}^k |\sigma_i| (\hat{C}_1 \max_{l-m \leq i \leq l} \epsilon_i + \hat{C}_2 \Delta^{\min\{p_2, p_3\}}) \leq \\ &\leq \epsilon_l + \Delta k \sigma \hat{C}_1 \max_{l-m \leq i \leq l} \epsilon_i + k \sigma \hat{C}_2 \Delta^{\min\{p_2, p_3\}+1} + C \Delta^{p_1+1}.\end{aligned}$$

Таким образом, имеем оценку

$$\epsilon_{l+1} \leq \epsilon_l + \Delta C_7 \max_{l-m \leq i \leq l} \epsilon_i + C_8 \Delta^{p+1}, \quad (3.24)$$

где $C_7 = k\sigma\hat{C}_1$, $C_8 = k\sigma\hat{C}_2 + C$, $p = \min\{p_1, p_2, p_3\}$.

Индукцией по l докажем оценку

$$\epsilon_l \leq (1 + \Delta(C_7 + 1))^l C_8 \Delta^p. \quad (3.25)$$

База индукции выполняется, так как $\epsilon_0 = 0$.

Шаг индукции. Пусть оценка (3.25) выполняется для индексов $\leq l$, покажем её справедливость для $l+1$.

Пусть \max в правой части оценки (3.24) достигается на индексе $i_0 \leq l$, тогда, применяя индуктивное предположение к ϵ_l и ϵ_{i_0} , получаем

$$\begin{aligned}\epsilon_{l+1} &\leq \epsilon_l + \Delta C_7 \epsilon_{i_0} + C_8 \Delta^p \leq \\ &\leq (1 + \Delta(C_7 + 1))^l C_8 \Delta^p + \\ &+ \Delta C_7 (1 + \Delta(C_7 + 1))^{i_0} C_8 \Delta^p + C_8 \Delta^p \leq \\ &\leq (1 + \Delta(C_7 + 1))^l C_8 \Delta^p (1 + C_7 \Delta + \Delta).\end{aligned}$$

Оценка (3.25) доказана.

Из этой оценки выводится (как в теореме 3.1) оценка

$$\epsilon_l \leq e^{(C_7+1)\theta} C_8 \Delta^p,$$

которая содержит утверждение теоремы. □

3.5. Порядок невязки ЯРК-методов

В обыкновенных дифференциальных уравнениях порядок невязки ЯРК-методов определяется разложением по формуле Тейлора точного решения и правой части системы.

Пример 3.1. Рассмотрим для системы ОДУ

$$\dot{x} = f(t, x)$$

метод Хойна, называемый также методом Эйлера с пересчетом:

$$u_{l+1} = u_l + \frac{\Delta}{2}(f(t_l, u_l) + f(t_l, u_l + \Delta f(t_l, u_l))).$$

Невязка этого метода определяется соотношением

$$\psi(t_l) = \frac{x_{l+1} - x_l}{\Delta} - \frac{1}{2}(f(t_l, x_l) + f(t_l, x_l + \Delta f(t_l, x_l))).$$

Покажем, что невязка метода Хойна имеет второй порядок на достаточно гладких решениях. Разлагая $x(t)$ по формуле Тейлора, имеем

$$\begin{aligned} x_{l+1} &= x_l + \dot{x}(t_l)\Delta + \ddot{x}(t_l)\frac{\Delta^2}{2} + O(\Delta^3) = \\ &= x_l + f(t_l, x_l)\Delta + \frac{\Delta^2}{2} \left[\frac{\partial f}{\partial t} + \frac{\partial f}{\partial x} f \right] + O(\Delta^3), \end{aligned}$$

где $O(\Delta^3)$ – величина третьего порядка малости по Δ ,

$$\frac{\partial f}{\partial x} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_l} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_l} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & \cdots & \frac{\partial f_n}{\partial x_l} \end{pmatrix}$$

— матрица Якоби в точке (t_l, x_l) .

Кроме того, из формулы Тейлора следует представление

$$f(t_l, x_l + \Delta f(t_l, x_l)) = f(t_l, x_l) + \left[\frac{\partial f}{\partial t} + \frac{\partial f}{\partial x} f \right] \Delta + O(\Delta^2).$$

Подставляя эти разложения в невязку, получаем $\psi(t_l) = O(\Delta^2)$.

Для функционально-дифференциальных уравнений методика определения порядка невязки с помощью разложения Тейлора сохраняется, однако возникает вопрос с вычислением производных векторного функционала правой части системы (3.1) $f(t, x, y(\cdot))$. Для решения этого вопроса рассмотрим элементы i -гладкого анализа. Отметим, что для получения тейлоровских разложений удобнее использовать коинвариантные производные $\partial_t f(t_l, x(t_l), x_{t_l}(\cdot))$.

В качестве примера рассмотрим для уравнения (3.1) формулу аналога метода Эйлера с пересчетом

$$u_0 = x^0,$$

$$u_{l+1} = u_l + \frac{\Delta}{2} \left(f(t_l, u_l, u_{t_l}(\cdot)) + f(t_l + \Delta, u_l + \Delta f(t_l, u_l, u_{t_l}(\cdot)), u_{t_l + \Delta}(\cdot)) \right),$$

где $u_t(\cdot)$ – результат действия некоторого оператора интерполяции и экстраполяции.

Опишем процедуру оценки порядка аппроксимации метода. Рассмотрим невязку метода

$$\psi(t_l) = \frac{x(t_{l+1}) - x(t_l)}{\Delta} - \frac{1}{2} \left(f(t_l, x(t_l), x_{t_l}(\cdot)) + f(t_l, x(t_l) + \Delta f(t_l, x(t_l), x_{t_l}(\cdot)), x_{t_{l+1}}(\cdot)) \right).$$

Разлагая точное решение в ряд Тейлора, получаем

$$\begin{aligned} x(t_{l+1}) &= x(t_l) + \dot{x}(t_l)\Delta + \ddot{x}(t_l)\frac{\Delta^2}{2} + O(\Delta^3) = x(t_k) + f(t_l, x(t_l), x_{t_l}(\cdot))\Delta + \\ &+ \frac{\Delta^2}{2} \left[\partial_t f(t_l, x(t_l), x_{t_l}(\cdot)) + \frac{\partial f(t_l, x(t_l), x_{t_l}(\cdot))}{\partial x} f(t_l, x(t_l), x_{t_l}(\cdot)) \right] + O(\Delta^3). \end{aligned}$$

Так же имеем

$$\begin{aligned} f(t_l + \Delta, x(t_l) + \Delta f(t_l, x(t_l), x_{t_l}(\cdot)), x_{t_{l+1}}(\cdot)) &= f(t_l, x(t_l), x_{t_l}(\cdot)) + \\ &+ \left[\partial_t f(t_l, x(t_l), x_{t_l}(\cdot)) + \frac{\partial f(t_l, x(t_l), x_{t_l}(\cdot))}{\partial x} f(t_l, x(t_l), x_{t_l}(\cdot)) \right] \Delta + O(\Delta^2). \end{aligned}$$

Подставляя эти формулы в невязку получаем $\psi(t_l) = O(\Delta^2)$.

Данный метод имеет второй порядок сходимости, если используется интерполяция и экстраполяция предыстории модели второго порядка.

Чтобы выписать слагаемые высших порядков, необходимо использовать соответствующие частные, инвариантные (или коинвариантные), а также смешанные производные высших порядков. Важно отметить, что правила, по которым считаются коэффициенты в тейлоровских разложениях точного решения и функционала правой части системы, такие же, что и для обыкновенных дифференциальных уравнений. Отсюда, при условии достаточной гладкости точного решения и функционала правой части системы, следует

Теорема 3.5. *Если некоторый ЯРК-метод имеет для ОДУ порядок невязки p , то ЯРК-метод с теми же коэффициентами для ФДУ имеет также порядок невязки p .*

Это утверждение, вместе с теоремой 3.4 о порядке сходимости, позволяет конструировать для ФДУ аналоги всех известных для ОДУ ЯРК-методов, выбирая, однако, подходящие способы интерполяции и экстраполяции предыстории модели. Так, аналог метода Эйлера с пересчетом с кусочно-линейной интерполяцией (3.11) и с экстраполяцией продолжением (3.17) имеет второй порядок сходимости.

Еще один *пример* — аналог известного четырехэтапного ЯРК-метода

$$\begin{aligned} u_{l+1} &= u_l + \frac{1}{6}\Delta (h_1 + 2h_2 + 2h_3 + h_4), \\ h_1 &= f(t_l, u_l, u_{t_l}(\cdot)), \quad h_2 = f(t_l + \frac{\Delta}{2}, u_l + \frac{\Delta}{2}h_1, u_{t_l + \frac{\Delta}{2}}(\cdot)), \\ h_3 &= f(t_l + \frac{\Delta}{2}, u_l + \frac{\Delta}{2}h_2, u_{t_l + \frac{\Delta}{2}}(\cdot)), \\ h_4 &= f(t_l + \Delta, u_l + \Delta h_3, u_{t_l + \Delta}(\cdot)). \end{aligned}$$

Этот метод имеет четвертый порядок сходимости (при условии достаточной гладкости решения), если производить интерполяцию предыстории модели кусочно-кубическими функциями и экстраполяцию – продолжением.

При порядке невязки $p \leq 5$ не существует ЯРК-метода с числом этапов $k = p$, этот факт называется *барьером Бутчера* [22].

Один из ЯРК-методов с числом этапов $k = 6$ порядка $p = 5$ – *метод Рунге–Кутты–Фельберга* положен в основу стандартных программ решения систем обыкновенных дифференциальных уравнений в таких пакетах, как MATLAB, MAPLE. Для ФДУ аналог метода с интерполяцией вырожденными сплайнами четвертой степени и экстраполяцией продолжением создан как инструментальный в системе MATLAB и он послужил основой для пакета Time-Delay System Toolbox [45]. Подробнее см. главу 4.

3.6. Неявные методы типа Рунге–Кутты

Неявные методы типа Рунге–Кутты по сравнению с явными обладают рядом достоинств:

- для любого числа этапов k можно построить метод с порядком сходимости $p = 2k$;
- лучше приспособлены для решения жестких систем;
- имеют преимущества при распараллеливании вычислений.

Пусть задан оператор интерполяции–экстраполяции IE .

В дальнейшем будем предполагать, что:

- 1) оператор IE *согласован*, т.е.

$$u(t_i) = u_i, \quad i = l - m, \dots, l,$$

- 2) оператор IE *удовлетворяет условию Липшица*, т.е. найдется константа L_I такая, что для всяких двух наборов дискретных предысторий $\{u_i^1\}_l$ и $\{u_i^2\}_l$ и для всех $t \in [t_l - \tau, t_l + a\Delta]$ выполняется

$$\|u^1(t) - u^2(t)\| \leq L_I \max_{l-m \leq i \leq l} \|u_i^1 - u_i^2\|,$$

где $u^1(\cdot) = IE(\{u_i^1\}_l)$, $u^2(\cdot) = IE(\{u_i^2\}_l)$.

Способы интерполяции и экстраполяции предыстории модели, указанные в разделе 3.1.2, удовлетворяют указанным условиям.

Определение 3.11. Для натурального k назовем k -этапным неявным методом типа Рунге–Кутты (НРК) (с интерполяцией I и экстраполяцией E) численную модель вида

$$u_0 = x_0; \tag{3.26}$$

$$\begin{aligned} u_{l+1} &= u_l + \Delta \Psi(t_l, u(\cdot)) = \\ &= u_l + \Delta \sum_{i=1}^k \sigma_i h_i(u_{t_l}(\cdot)), \quad l = 1, \dots, N-1, \end{aligned} \tag{3.27}$$

$$h_i(u_{t_l}(\cdot)) = f(t_l + a_i \Delta, u_l + \Delta \sum_{j=1}^k b_{ij} h_j(u_{t_l}(\cdot)), u_{t_l + a_i \Delta}(\cdot)). \tag{3.28}$$

Здесь предыстория модели определяется соотношением $u(\cdot) = IE(\{u_i^1\}_l)$.

Коэффициенты метода удобно задавать таблицей Бутчера

$$\begin{array}{c|cccc} a_1 & b_{11} & b_{12} & \dots & b_{1k} \\ a_2 & b_{21} & b_{22} & \dots & b_{2k} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ a_k & b_{k1} & b_{k2} & \dots & b_{kk} \\ \hline & \sigma_1 & \sigma_2 & \dots & \sigma_k \end{array}$$

Отметим некоторые частные случаи:

- 1) если $b_{ij} = 0$ при $j \geq i$, то получаем ЯРК-метод;
- 2) если $b_{ij} = 0$ при $j > i$, то метод называется диагонально-неявным;
- 3) если $b_{ij} = 0$ при $j > i$ и $b_{ii} = b_0$, то метод называется однократно-диагональным.

Как и в разделе 3.1.3 будем обозначать $a = \max\{|a_i|, 1 \leq i \leq k\}$, $\sigma = \max\{|\sigma_i|\}$, $b = \max\{|b_{ij}|\}$.

Теорема 3.6. Пусть

$$\Delta < \frac{1}{kbL}.$$

Тогда существует единственное решение $J = (h_1, h_2, \dots, h_k)'$ системы уравнений (3.28), причем функционалы h_1, h_2, \dots, h_k , а также Ψ липшицевы по $u(\cdot)$.

Доказательство. В векторном пространстве \mathbb{R}^{kn} для вектора $J = (h_1, h_2, \dots, h_k)'$ введем норму $\|J\| = \max_i \|h_i\|$.

Систему уравнений (3.28) можно записать в виде

$$J = \Phi(J).$$

Покажем, что при малом шаге оператор Φ сжимающий. Пусть $J^1 = (h_1^1, h_2^1, \dots, h_k^1)'$, $J^2 = (h_1^2, h_2^2, \dots, h_k^2)'$, тогда

$$\begin{aligned} & \|\Phi(J^1) - \Phi(J^2)\| = \\ &= \max_i \left\| f(t_l + a_i \Delta, u_l + \Delta \sum_{j=1}^k b_{ij} h_j^1(u_{t_l}(\cdot)), u_{t_l + a_i \Delta}(\cdot)) - \right. \\ & \quad \left. - f(t_l + a_i \Delta, u_l + \Delta \sum_{j=1}^k b_{ij} h_j^2(u_{t_l}(\cdot)), u_{t_l + a_i \Delta}(\cdot)) \right\| = \\ &= \left\| f(t_l + a_{i_0} \Delta, u_l + \Delta \sum_{j=1}^k b_{i_0 j} h_j^1(u_{t_l}(\cdot)), u_{t_l + a_{i_0} \Delta}(\cdot)) - \right. \\ & \quad \left. - f(t_l + a_{i_0} \Delta, u_l + \Delta \sum_{j=1}^k b_{i_0 j} h_j^2(u_{t_l}(\cdot)), u_{t_l + a_{i_0} \Delta}(\cdot)) \right\| \leq \\ & \leq L \Delta b \sum_{j=1}^k \|h_j^1 - h_j^2\| \leq L \Delta b k \|J^1 - J^2\| \end{aligned}$$

и при сделанных предположениях на шаг

$$\|\Phi(J^1) - \Phi(J^2)\| \leq \gamma \|J^1 - J^2\|, \quad 0 < \gamma < 1.$$

Применяя принцип сжатых отображений к итерационной процедуре $J^{i+1} = \Phi(J^i)$, получаем в пределе решение системы (3.28).

Докажем липшицевость h_i . Пусть функции $u^1(\cdot) \in Q[t_l - \tau, t_l + a\Delta]$, $u^2(\cdot) \in Q[t_l - \tau, t_l + a\Delta]$, $u_l^1 = u^1(t_l)$, $u_l^2 = u^2(t_l)$, $h_i^1 = f(t_l + a_i\Delta, u_l^1 + \Delta \sum_{j=1}^k b_{ij} h_j^1, u_{t_l+a_i\Delta}^1(\cdot))$, $h_i^2 = f(t_l + a_i\Delta, u_l^2 + \Delta \sum_{j=1}^k b_{ij} h_j^2, u_{t_l+a_i\Delta}^2(\cdot))$, тогда

$$\begin{aligned} \|h_i^1 - h_i^2\| &\leq L(\|u_l^1 - u_l^2\| + b\Delta \sum_{j=1}^k \|h_j^1 - h_j^2\|) + \\ &+ M \sup_{-\tau \leq s < 0} \|u_{t_l+a_i\Delta}^1(s) - u_{t_l+a_i\Delta}^2(s)\| \leq \\ &\leq b\Delta L \sum_{j=1}^k \|h_j^1 - h_j^2\| + (L + M) \sup_{t_l - \tau \leq t < t_l + a\Delta} \|u^1(t) - u^2(t)\|. \end{aligned}$$

Для разности $J^1 = (h_1^1, h_2^1, \dots, h_k^1)'$ и $J^2 = (h_1^2, h_2^2, \dots, h_k^2)'$ справедливо

$$\|J^1 - J^2\| \leq bk\Delta L \|J^1 - J^2\| + (L + M) \sup_{t_l - \tau \leq t < t_l + a\Delta} \|u^1(t) - u^2(t)\|$$

или

$$\|h_i^1 - h_i^2\| \leq \|J^1 - J^2\| \leq \frac{L + M}{1 - bk\Delta L} \sup_{t_l - \tau \leq t < t_l + a\Delta} \|u^1(t) - u^2(t)\|.$$

Константа

$$L_J = \frac{L + M}{1 - bk\Delta L}$$

положительна в силу предположения теоремы.

Из липшицевости h_i и определения функции Ψ (3.28) вытекает липшицевость Ψ по $u(\cdot)$ с константой Липшица $L_\Psi = L_J k a$, т.е.

$$\|\Psi(t_l, u^1(\cdot)) - \Psi(t_l, u^2(\cdot))\| \leq L_\Psi \sup_{t_l - \tau \leq t < t_l + a\Delta} \|u^1(t) - u^2(t)\|.$$

□

Следствие 3.1. Если оператор интерполяции-экстраполяции IE задан, то функция $\Phi(\{u_i\}_l) = \Psi(t_l, IE(\{u_i\}_l))$ липшицева по дискретной предыстории модели, т.е. найдется такая константа $L_\Phi = L_\Psi L_I$, что выполняется неравенство

$$\|\Phi(\{u_i^1\}_l) - \Phi(\{u_i^2\}_l)\| \leq L_\Phi \max_{l-m \leq i \leq l} \|u_i^1 - u_i^2\|.$$

Порядок невязки и порядок сходимости НРК-методов определяются так же, как и ЯРК-методов.

Теорема 3.7. Если метод (3.26) - (3.28) имеет невязку порядка $p_1 > 0$, оператор интерполяции-экстраполяции предыстории модели имеет порядок $p_2 > 0$, то метод сходится, причем порядок сходимости НРК-метода p не меньше минимума из p_1, p_2 .

Эта теорема доказывается подобно теореме 3.5.

Замечание 3.1. Для ФДУ, так же, как и для ОДУ, для любого числа этапов k можно построить метод с порядком сходимости $p = 2k$.

При $k = 1$ таким методом является неявное правило средней точки: НРК-метод с таблицей Бутчера

$$\begin{array}{c|c} \frac{1}{2} & \frac{1}{2} \\ \hline & 1 \end{array}.$$

При $k = 2$ таким является метод Хаммера-Холлингсуорта порядка 4:

$$\begin{array}{cc|cc} \frac{1}{2} - \frac{\sqrt{3}}{6} & & \frac{1}{4} & \frac{1}{4} - \frac{\sqrt{3}}{6} \\ \frac{1}{2} + \frac{\sqrt{3}}{6} & & \frac{1}{4} + \frac{\sqrt{3}}{6} & \frac{1}{4} \\ \hline & & \frac{1}{2} & \frac{1}{2} \end{array}.$$

Замечание 3.2. При реализации НРК-методов на каждом шаге приходится решать систему уравнений (3.27) относительно u_{l+1} (или систему (3.28) относительно h_i). При этом возможны следующие реализации.

1) Схема предиктор-корректор. Делают прогноз u_{l+1}^0 каким-то явным методом, а затем одну итерацию по методу простой итерации

$$u_{l+1}^1 = u_l^1 + \Delta \sum_{i=1}^k \sigma_i h_i(u_{t_i}^0(\cdot)).$$

2) Схема $P(EC)^m$, т.е. предиктор делается один раз, а вычисление-корректор — m раз:

$$u_{l+1}^{i+1} = u_l^m + \Delta \sum_{i=1}^k \sigma_i h_i(u_{t_i}^i(\cdot)), \quad i = 0, \dots, m-1.$$

Вариант этой схемы - схема $P(EC)^m E$, в которой значение функционала правой части системы вычисляется еще один раз для использования на последующем шаге.

3) Градиентный метод Ньютона-Рафсона для решения нелинейной системы. Основная вычислительная проблема этого метода — счет матрицы Якоби.

Замечание 3.3. НРК-методы для ОДУ в настоящее время являются основным способом решения жестких систем [25, 23]. Для ФДУ этот вопрос пока недостаточно изучен.

Замечание 3.4. Возможности для параллельных вычислений. Неявные методы Рунге–Кутты для ОДУ обладают значительными возможностями для организации параллельных вычислений. В работе [41] отмечены три основных направления распараллеливания, которые сохраняются и для ФДУ.

1. Использование НРК-методов с матрицей Бутчера такого вида, который позволяет применять параллельно несколько процессоров для вычисления величин h_i по формулам (3.28). Например, четырехэтапный метод

$$\begin{array}{c|cccc} \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ \frac{2}{3} & 0 & \frac{2}{3} & 0 & 0 \\ \frac{1}{2} & -\frac{5}{2} & \frac{5}{2} & \frac{1}{2} & 0 \\ \frac{1}{3} & -\frac{5}{3} & \frac{4}{3} & 0 & \frac{2}{3} \\ \hline & -1 & \frac{3}{2} & -1 & \frac{3}{2} \end{array}$$

позволяет использовать параллельно два процессора для вычисления h_1 и h_2 , а затем h_3 и h_4 . Более содержательные примеры приведены в [41], причем там отмечается, что преимущества параллельных вычислений проявляются именно для неявных методов Рунге–Кутты.

2. Чтобы вычислить h_i в (3.28) необходимо решить нелинейные системы. Обычно для этого применяют метод Ньютона–Рафсона, на каждой итерации которого требуется решать линейные системы нередко большой размерности. Методы факторизации матрицы Якоби, применяемые в ОДУ для распараллеливания этих вычислений, применимы и для ФДУ.

3. Как для обыкновенных дифференциальных уравнений, так и для ФДУ имеется возможность распараллеливания вычисления координат векторной функции $f(t_i, u_i, u_{t_i}(\cdot))$ — правой части системы (6.1), что особенно актуально, если размерность n системы велика.

В ФДУ для алгоритмов НРК имеется еще несколько возможностей для распараллеливания.

4. Интерполяция и экстраполяция предыстории дискретной модели проводятся параллельно по всем координатам, что существенно при $n > 1$.

5. Во всех конструкциях НРК-методов основные вычислительные ресурсы тратятся на подсчет функционала $f(t, x(t), x_t(\cdot))$ правой части системы (6.1), особенно если этот функционал содержит *распределенное* последствие. Пусть функционал f имеет вид

$$f = F(t, q_1, q_2, \dots, q_s), \quad (3.29)$$

где $F : \mathbb{R}^{s+1} \rightarrow \mathbb{R}^n$, $q_i = q_i(t, x(t), x_t(\cdot))$, $i = 1, \dots, s$, — сингулярные или регулярные (см. главу 1) функционалы над $\mathbb{R} \times \mathbb{R}^n \times Q[-\tau, 0]$.

Основные проблемы при этом возникают в том случае, если q_i — это интегралы вида

$$q_i = \int_{-\tau}^0 \beta_i(s, x(t), x_t(s)) ds \quad (3.30)$$

или имеют еще более сложную структуру см. выше, главы 1 и 2. В этом случае вычисления q_i , $i = 1, \dots, s$, можно проводить параллельно. В частности, к виду (3.29) —

(3.30) сводятся интегродифференциальные уравнения Вольтерра [25], рассматриваемые на конечном промежутке времени.

6. Если функционал f имеет распределенное запаздывание в виде интегралов (3.30) и подынтегральные функции $\beta_i(s, x(t), x_t(s))$ имеют сложную структуру или их вычисления трудоемки, то для численного вычисления интегралов (3.30) можно использовать, например, составные квадратурные формулы с параллельным вычислением значений подынтегральных функций.

Глава 4.

Многошаговые и некоторые другие численные методы для ФДУ

4.1. Многошаговые методы

Наряду с методами типа Рунге–Кутты при решении ОДУ самыми распространенными являются многошаговые линейные методы, в частности, методы типа Адамса.

Пусть так же, как и в методах типа Рунге–Кутты заданы временная сетка и оператор интерполяции-экстраполяции предыстории модели $IE : \{u_i\}_l \rightarrow u(\cdot) \in Q[t_l - \tau, t_l + \Delta]$.

Определение 4.1. *Явным k -шаговым методом назовем дискретную модель вида*

$$u_l = \sum_{i=1}^k \alpha_i u_{l-i} + \Delta \sum_{i=1}^k \beta_i f_{l-i}^*, \quad l = k, \dots, N, \quad (4.1)$$

где α_i и β_i ($i = 1, \dots, k$) – параметры метода, $f_i^* = f(t_i, u_i, u_{t_i}(\cdot))$.

Определение 4.2. *Неявным k -шаговым методом назовем модель*

$$u_l = \sum_{i=1}^k \alpha_i u_{l-i} + \Delta \sum_{i=0}^k \beta_i f_{l-i}^*, \quad l = k, \dots, N. \quad (4.2)$$

Отметим, что для применения моделей (4.1) или (4.2) при $l = k, \dots$ необходимо знание *стартовых* значений u_i , $i = 0, \dots, k-1$, которые определяют f_i^* , $i = 0, \dots, k-1$ (*разгон*).

Определение 4.3. *Будем говорить, что задан разгон порядка p , если найдется $C > 0$ такое, что $\|x_i - u_i\| \leq C \Delta^p$ для $i = 0, \dots, k-1$.*

Можно также отметить, что экстраполяция нужна лишь для неявных методов.

Для оценки погрешности метода введем величину $z_l = u_l - x_l$ и подставим $u_{l-i} = x_{l-i} + z_{l-i}$, $i = 0, \dots, k$, в (4.2). Получим

$$z_l = \sum_{i=1}^k \alpha_i z_{l-i} + \Delta \psi_l^{(1)} + \Delta \psi_l^{(2)},$$

где

$$\begin{aligned} \psi_l^{(1)} &= \frac{1}{\Delta} \left(\sum_{i=1}^k \alpha_i x_{l-i} - x_l \right) + \sum_{i=0}^k \beta_i f_{l-i}, \\ f_i &= f(t_i, x(t_i), x_{t_i}(\cdot)), \\ \psi_l^{(2)} &= \sum_{i=0}^k \beta_i (f_{l-i}^* - f_{l-i}). \end{aligned} \quad (4.3)$$

Определение 4.4. Величину $\psi_l^{(1)}$ будем называть погрешностью аппроксимации многошагового метода (невязкой) и будем говорить, что невязка имеет порядок p , если найдется C такое, что $\|\psi_l^{(1)}\| \leq C\Delta^p$ для всех $l = k, \dots, N$.

Определение 4.5. Будем говорить, что метод (4.2) сходится, если $\|z_l\| \rightarrow 0$ при $\Delta \rightarrow 0$ для всех $l = 0, \dots, N$, и имеет порядок сходимости p , если найдется C такое, что $\|z_l\| \leq C\Delta^p$ для всех $l = k, \dots, N$.

Определение 4.6. Будем говорить, что метод (4.2) 0-устойчив, если выполняется условие корней, т.е. все корни ξ производящего многочлена

$$\rho^k - \sum_{i=1}^k \alpha_i \rho^{l-i}$$

удовлетворяют условию $|\xi| \leq 1$, причем если $|\xi| = 1$, то корень простой.

Теорема 4.1. Пусть многошаговый метод (4.2) 0-устойчив и имеет:

- 1) порядок погрешности аппроксимации p_1 ,
 - 2) разгон порядка p_2 ,
 - 3) порядок интерполяции и экстраполяции предыстории модели p_3 .
- Тогда он сходится, причем порядок сходимости $p \geq \min\{p_1, p_2, p_3\}$.

Теорема доказывается путём сведения к одношаговому методу за счет повышения размерности модели, см. для ОДУ [?, 19, 22].

Доказательство также следует из теоремы 5.1 и вложения многошаговых методов в общую схему из раздела 3.2.

Рассмотрим погрешность аппроксимации многошаговых методов.

Пусть точное решение $x(t)$ задачи (3.1) - (3.3) непрерывно дифференцируемо $p+1$ раз на $[t_0, t_0 + \theta]$. Из формулы Тейлора имеем

$$x_{l-i} = \sum_{j=0}^p \frac{(-i\Delta)^j x^{(j)}(t_l)}{j!} + O(\Delta^{p+1}),$$

$$f_{l-i} = \sum_{j=0}^p \frac{(-i\Delta)^j x^{(j+1)}(t_l)}{j!} + O(\Delta^p).$$

Подставив эти выражения в невязку (4.3), получим, что порядок невязки метода (4.2) не меньше $p \geq 1$, если выполняются условия

$$\sum_{i=1}^k \alpha_i = 1,$$

$$\sum_{i=0}^k i^{j-1} (j \beta_i - i \alpha_i) = 0, \quad j = 1, \dots, p.$$

Отсюда следует, что порядок аппроксимации для неявных методов $p \leq 2k$, для явных $p \leq 2k - 1$.

Однако как и в случае ОДУ для 0-устойчивых k -шаговых методов выполняются барьеры Далквиста[22]:

- 1) $p \leq k + 2$ для четных k ,
- 2) $p \leq k + 1$ для нечетных k ,
- 3) $p \leq k$, если $\beta_0 \leq 0$ (в частности, для явных методов).

Для методов Адамса

$$u_l = u_{l-1} + \Delta \left(\sum_{i=0}^k \beta_i f_{l-i}^* \right)$$

порядок погрешности аппроксимации $p \leq k + 1$ для неявных методов и $p \leq k$ для явных. Поэтому, по сравнению с методами Адамса, в общем классе многошаговых методов можно добиться большей точности только для неявных методов и только для чётных k .

Замечание 4.1. При реализации неявных методов на каждом шаге приходится решать систему уравнений (4.2) относительно u_l . При этом возможны следующие реализации.

1) *Схема предиктор-корректор.* Делают прогноз u_l^0 каким-то явным методом, а затем одну итерацию по методу простой итерации

$$u_l^1 = \sum_{i=1}^k \alpha_i u_{l-i} + \Delta \sum_{i=1}^k \beta_i f_{l-i}^* + \Delta \beta_0 f(t_l, u_l^0, u_{t_l}(\cdot)).$$

2) *Схема $P(EC)^m$,* т.е. предиктор делается один раз, а вычисление-корректор делают m раз:

$$u_l^{i+1} = \sum_{i=1}^k \alpha_i u_{l-i} + \Delta \sum_{i=1}^k \beta_i f_{l-i}^* + \Delta \beta_0 f(t_l, u_l^i, u_{t_l}(\cdot)), \quad i = 0, \dots, m-1.$$

Вариант этой схемы — *схема $P(EC)^m E$* , в которой значение функционала правой части системы вычисляется еще один раз для использования на последующем шаге.

3) *Градиентный метод* Ньютона–Рафсона для решения нелинейной системы (4.2). Основная вычислительная проблема этого метода — счет матрицы Якоби.

4.2. Многошаговые методы, не требующие разгона

Специфика ФДУ состоит в том, что изначально известны из (3.2) - (3.3) значения модели u_i при $-m \leq i \leq 0$. Однако применять формулы (4.1) или (4.2) без разгона, т.е. при $l = 1, 2, \dots, k-1$, вообще говоря, нельзя, так как неопределены значения f_i^* при $i < 0$.

Определение 4.7. Если $m \geq k$, то явным k -шаговым методом, не требующим разгона, назовем модель

$$u_l = \sum_{i=1}^k \alpha_i u_{l-i} + \Delta \beta_1 f_{l-1}^*, \quad l = 1, 2, \dots, N, \quad (4.4)$$

где

$$u_i = x_i \equiv y^0(t_i - t_0) \quad \text{при} \quad -m \leq i < 0, \\ u_0 = x_0.$$

Как следует из предыдущего раздела, порядок невязки p метода (4.4) удовлетворяет условию $p \leq k$. 0-устойчивыми методами (4.4), для которых $p = k$, являются: явный метод Эйлера

$$u_l = u_{l-1} + \Delta f_{l-1}^*, \quad p = k = 1,$$

и правило средней точки

$$u_l = u_{l-2} + 2 \Delta f_{l-1}^*, \quad p = k = 2.$$

При $k \geq 3$ среди методов вида (4.4), для которых $p = k$, нет 0-устойчивых. Однако существуют явные четырехшаговые 0-устойчивые методы третьего порядка аппроксимации, не требующие разгона, например,

$$u_l = -\frac{2}{5} u_{l-1} + \frac{6}{5} u_{l-2} + \frac{2}{5} u_{l-3} - \frac{1}{5} u_{l-4} + \frac{12}{5} \Delta f_{l-1}^*,$$

для данного метода $k = 4, p = 3$.

Определение 4.8. При $m \geq k$ неявным k -шаговым методом, не требующим разгона, назовем модель

$$u_l = \sum_{i=1}^k \alpha_i u_{l-i} + \Delta (\beta_0 f_l^* + \beta_1 f_{l-1}^*), \quad l = 1, \dots, N. \quad (4.5)$$

Порядок невязки p метода (4.5) удовлетворяет условию $p \leq k + 1$.

Среди методов вида (4.5), для которых $p = k + 1$, 0-устойчивым является лишь метод трапеций

$$u_l = u_{l-1} + \frac{\Delta}{2} (f_l^* + f_{l-1}^*), \quad k = 1, p = 2.$$

Среди методов вида (4.5), для которых $p = k$, отметим *формулы дифференцирования назад* (ФДН)

$$\sum_{i=1}^k \frac{1}{i} \nabla^i u_l = \Delta f_l^*$$

(здесь ∇ означает разность назад: $\nabla^0 u_l = u_l$, $\nabla^{i+1} u_l = \nabla^i u_l - \nabla^i u_{l-1}$), 0-устойчивые при $k \leq 6$ [22]. Эти методы применяются при решении жестких систем.

Например, двухшаговый метод второго порядка имеет вид

$$\frac{3u_l - 4u_{l-1} + u_{l-2}}{2} = \Delta f_l^*.$$

4.3. Методы Нордсика

Методы Нордсика [22] для ОДУ являются одними из самых экономичных. Они эквивалентны многошаговым, но в них легко можно изменять величину шага. Коэффициенты вектора Нордсика приближают производные точного решения системы, но сами производные считаются только один раз для получения стартовых значений. В случае ФДУ для счета производных эффективно использовать i -гладкий анализ.

В настоящем разделе для простоты обозначений будем рассматривать только одномерные системы (3.1), т.е. в которых $x \in \mathbb{R}$.

Пусть так же, как и в предыдущих разделах заданы временная сетка и оператор интерполяции-экстраполяции предыстории модели $IE : \{u_i\}_l \rightarrow u(\cdot) \in Q[t_l - \tau, t_l + \Delta]$.

Введем вектор $U_l \in \mathbb{R}^k$ (вектор Нордсика) с координатами $U_l = (u_l^1, u_l^2, \dots, u_l^k)$, первая координата которого $u_l^1 = u_l$ приближает точное решение $x(t_l)$, вторая u_l^2 приближает $\Delta \dot{x}(t_l)$, и так далее, последняя координата u_l^k приближает $\frac{\Delta^{k-1}}{(k-1)!} x^{(k-1)}(t_l)$. Здесь $l = 0, \dots, N$.

Методом Нордсика назовем пошаговую модель вида

$$U_{l+1} = PU_l + q[\Delta f(t_{l+1}, u_{l+1}, u_{t_{l+1}}(\cdot)) - e_2' P u_l], \quad l = 0, \dots, N-1, \quad (4.6)$$

где P – треугольная матрица Паскаля размерности $k \times k$, элементы которой P_{ij} определяются формулой:

$$P_{ij} = \begin{cases} C_j^i, & \text{при } 1 \leq i \leq j \leq k, \\ 0, & \text{в противном случае.} \end{cases}$$

Здесь C_j^i – число сочетаний из j по i ; вектор e_2 определяется соотношением $e_2 = (0, 1, 0, \dots, 0)'$; координаты вектора $q = (q_1, 1, q_3, \dots, q_l)'$ являются параметрами вектора.

Условие $q_2 = 1$ дает связь метода с уравнением (3.1): из пошаговой формулы вытекает $u_{l+1}^2 = f(t_{l+1}, u_{l+1}, u_{t_{l+1}}(\cdot))$.

Введем матрицу $S = P - qe_2'P$, тогда пошаговую формулу можно переписать в виде

$$U_{l+1} = SU_l + q\Delta f(t_{l+1}, u_{l+1}, u_{t_{l+1}}(\cdot)), \quad l = 0, \dots, N-1. \quad (4.7)$$

Непосредственно проверяется, что 1 и 0 являются собственными значениями матрицы S , а ее характеристический многочлен не зависит от параметра q_1 . Можно подобрать [?] числа q_3, \dots, q_k таким образом, чтобы остальные собственные значения были равны нулю. В этом случае метод 0-устойчив в смысле определения раздела 3.1.6. Коэффициент q_1 можно выбрать из условия обращения в нуль константы погрешности метода [22, с. 384]. Например, в случае $k = 4$ матрица S имеет вид

$$S = \begin{pmatrix} 1 & 1 - q_1 & 1 - 2q_1 & 1 - 3q_1 \\ 0 & 0 & 0 & 0 \\ 0 & -q_3 & 1 - 2q_3 & 3 - 3q_3 \\ 0 & -q_4 & -2q_4 & 1 - 3q_3 \end{pmatrix},$$

а вектор $q = \left(\frac{3}{8}, 1, \frac{3}{4}, \frac{1}{6}\right)^T$.

Стартовые значения метода определяются по формулам

$$\begin{aligned} u_0^1 &= x(t_0) = x_0, \\ u_0^2 &= \Delta \dot{x}(t_0) = \Delta f(t_0, x_0, y^0(\cdot)), \\ &\vdots \\ u_0^k &= \frac{\Delta^{k-1}}{(k-1)!} x^{(k-1)}(t_0). \end{aligned}$$

При вычислении производных функционала правой части уравнения (3.1) может быть использована техника i -гладкого анализа. Так,

$$\begin{aligned} u_0^3 &= \frac{\Delta^2}{2} \frac{d}{dt} [f(t, x, y(\cdot))] |_{t=t_0, x=x_0, y=y^0} = \\ &= \frac{\Delta^2}{2} [\partial_t f(t_0, x_0, y^0(\cdot)) + \frac{\partial f(t_0, x_0, y^0(\cdot))}{\partial x} f(t_0, x_0, y^0(\cdot))]. \end{aligned}$$

Определим дискретную функцию точных значений – вектор $Z_l \in \mathbb{R}^k$ с координатами $Z_l = (z_l^1, z_l^2, \dots, z_l^k)$ соотношениями

$$z_l^1 = x(t_l), \quad z_l^j = \frac{\Delta^{j-1}}{(j-1)!} x^{(j-1)}(t_l), \quad j = 2, \dots, k, \quad l = 0, \dots, N.$$

Невязкой (погрешностью аппроксимации) метода Нордсика назовем функцию

$$\psi(t_l) = \frac{Z_{l+1} - SZ_l}{\Delta} - qf(t_{l+1}, x_{l+1}, x_{t_{l+1}}(\cdot)), \quad l = 0, \dots, N-1.$$

Будем говорить, что невязка имеет порядок p , если найдется постоянная C такая, что $\|\psi(t_l)\| \leq C\Delta^p$ для всех $l = 0, 1, \dots, N-1$.

Будем говорить, что метод (4.7) *сходится*, если $\|Z_l - U_l\| \rightarrow 0$ при $\Delta \rightarrow 0$ для всех $l = 0, \dots, N$, и имеет *порядок сходимости* p , если найдется C такое, что $\|Z_l - U_l\| \leq C\Delta^p$ для всех $l = k, \dots, N$.

Теорема 4.2. Пусть в методе Нордсика (4.7) координаты вектора q выбраны так, чтобы обеспечить 0-устойчивость, кроме того метод имеет невязку порядка p_1 и порядок интерполяции и экстраполяции предыстории модели равен p_2 . Тогда он сходится, причем порядок сходимости $p \geq \min\{p_1, p_2\}$.

Доказательство следует из утверждения [22, с. 382]: метод Нордсика эквивалентен многошаговым методам, а также следует из теоремы 3.1 и вложения метода Нордсика в общую схему раздела 3.2.

Замечание 4.2. В многошаговых методах изменить порядок метода просто: надо только добавить следующее слагаемое в представлении (4.2). Изменение шага является достаточно сложной процедурой. В представлении Нордсика (4.6), наоборот, изменение порядка не является столь же естественным, однако можно легко изменить шаг, умножив вектор Нордсика на диагональную матрицу с элементами по диагонали $(1, \omega, \omega^2, \dots)$, где $\omega = h_{new}/h_{old}$ – отношение шагов.

4.4. Методы, использующие вычисление старших производных

Если задано ФДУ (3.1) то, используя тейлоровское разложение точного решения $x(t)$ или функционала f , можно выписать метод любого порядка точности. Например, формула

$$x(t_{l+1}) \approx x(t_l) + \Delta \dot{x}(t_l) + \frac{\Delta^2}{2} \ddot{x}(t_l)$$

при использовании правил вычисления производных функционала вдоль решения

$$\begin{aligned}\ddot{x}(t) &= \frac{d}{dt}f(t, x(t), x_t(\cdot)) = \\ &= \partial_t f(t, x(t), x_t(\cdot)) + \frac{\partial f(t, x(t), x_t(\cdot))}{\partial x} f(t, x(t), x_t(\cdot))\end{aligned}$$

даёт (в сочетании с выбранным оператором интерполяции предыстории модели $I : \{u_i\}_I \rightarrow u(\cdot) \in Q[t_l - \tau, t_l]$) метод

$$\begin{aligned}u_{l+1} &= u_l + \Delta f(t_l, u_l, u_{t_l}(\cdot)) + \\ &+ \frac{\Delta^2}{2} \left[\partial_t f(t_l, u_l, u_{t_l}(\cdot)) + \frac{\partial f(t_l, u_l, u_{t_l}(\cdot))}{\partial x} f(t_l, u_l, u_{t_l}(\cdot)) \right].\end{aligned}\quad (4.8)$$

Заметим, что в отличие от методов типа Рунге–Кутты, экстраполяция дискретной предыстории модели не нужна. Этот метод имеет второй порядок невязки, что непосредственно следует из тейлоровского разложения, и потому рационально применять кусочно-линейную интерполяцию. Порядок сходимости в этом случае равен двум, это утверждение доказывается подобно теореме 3.4.

Такие методы удобно применять в случае, когда структура функционала правой части системы (3.1) фиксирована.

Пример 4.1. Рассмотрим линейную систему с постоянным сосредоточенным и распределенным запаздыванием

$$\dot{x}(t) = Ax(t) + Bx(t - \tau) + \int_{-\tau}^0 G(s)x(t + s)ds. \quad (4.9)$$

Такие системы часто возникают в теории управления и стабилизации систем с запаздыванием. При этом, даже если исходный управляемый объект имеет лишь сосредоточенное запаздывание

$$\dot{x}(t) = A_1x(t) + Bx(t - \tau) + Cv(t),$$

наличие управления с обратной связью

$$v(t) = Px(t) + \int_{-\tau}^0 Q(s)x(t + s)ds$$

приводит к системе с распределенным запаздыванием вида (4.9), аналитическое решение которой затруднено.

Используя технику i -гладкого анализа, вычислим градиент и коинвариантную производную (она в данном случае совпадает с инвариантной производной) функционала $f(t, x, y(\cdot)) = Ax + By(-\tau) + \int_{-\tau}^0 G(s)y(s)ds$ — правой части системы (4.9).

$$\frac{\partial f}{\partial x} = A;$$

$$\partial_t f = B\dot{x}(t - \tau) + G(0)x(t) - G(-\tau)x(t - \tau).$$

Если обозначим (как и в предыдущих разделах $l = 0, \dots, N$, $N = \theta/\Delta$, $m = \tau/\Delta$)

$$f_l = f(t_l, u_l, u_{t_l}(\cdot)) = Au_l + Bu_{l-m} + \int_{-\tau}^0 G(s)u(t_l + s)ds, \quad (4.10)$$

то метод (4.8) примет вид

$$u_{l+1} = u_l + \Delta f_l + \frac{\Delta^2}{2} [Bf_{l-m} + G(0)u_l - G(-\tau)u_{l-m} + Af_l]. \quad (4.11)$$

Аналогичным образом можно использовать производные более высоких порядков. Систему (3.1) покоординатно можно записать

$$\dot{x}_i(t) = f_i(t, x(t), x_t(\cdot)), \quad i = 1, \dots, n,$$

тогда

$$\begin{aligned} \frac{d^2}{dt^2}(x_i(t)) &= \partial_t f_i(t, x(t), x_t(\cdot)) + \frac{\partial f_i(t, x(t), x_t(\cdot))}{\partial x_1} f_1(t, x(t), x_t(\cdot)) + \\ &+ \dots + \frac{\partial f_i(t, x(t), x_t(\cdot))}{\partial x_n} f_n(t, x(t), x_t(\cdot)). \end{aligned} \quad (4.12)$$

Введем обозначения. Пусть функционал $\Psi(t, x(t), x_t(\cdot)) : [t_0, t_0 + \theta] \times \mathbb{R}^n \times Q[-\tau, 0] \rightarrow \mathbb{R}$ обладает достаточными для дальнейшего свойствами гладкости. Обозначим через D дифференциальный оператор, действующий на функционал Ψ по правилу

$$\begin{aligned} (D\Psi)(t, x(t), x_t(\cdot)) &= \partial_t \Psi(t, x(t), x_t(\cdot)) + \frac{\partial \Psi(t, x(t), x_t(\cdot))}{\partial x_1} f_1(t, x(t), x_t(\cdot)) + \\ &+ \dots + \frac{\partial \Psi(t, x(t), x_t(\cdot))}{\partial x_n} f_n(t, x(t), x_t(\cdot)). \end{aligned}$$

Поскольку $Dx_i = f_i$, то получаем формулы, обобщающие (4.12):

$$\frac{d^j}{dt^j}(x_i(t)) = (D^j x_i)(t, x(t), x_t(\cdot)).$$

Эти обозначения удобны для определения нового класса численных методов решения ФДУ, сочетающих в себе черты как методов типа Рунге–Кутты, так и методов, использующих тейлоровское разложение.

Для натурального k назовем k -этапным явным методом типа Рунге–Кутты с использованием производных порядка q (и с оператором интерполяции-экстраполяции IE) численную модель вида

$$u_{l+1} = u_l + \sum_{r=1}^q \sum_{j=1}^k \sigma_j^r h_j^r(u_l, u_{t_l}(\cdot)), \quad l = 1, \dots, N-1, \quad (4.13)$$

$$h_i^n(u_l, u_{t_l}(\cdot)) = \frac{\Delta^n}{n!} (D^n x)(t_l + a_i \Delta, u_l + \sum_{r=1}^q \sum_{j=1}^k b_{ij}^r h_j^r(u_l, u_{t_l}(\cdot)), u_{t_l + a_i \Delta}(\cdot)).$$

Здесь предыстория модели $u_t(\cdot)$ определяется оператором IE .

Числа

$$\begin{aligned} a_i, \quad i = 1, \dots, k, \\ \sigma_j^r, \quad j = 1, \dots, k, \quad r = 1, \dots, q, \\ b_{ij}^r, \quad i, j = 1, \dots, k, \quad r = 1, \dots, q \end{aligned}$$

— коэффициенты метода, который является явным, если все b_{ij}^r равны нулю при $i \leq j$, и неявным — в противном случае.

Эти методы эквивалентны *коллокационным* с кратными узлами [22, с. 267], и содержат такие важные классы, как методы Обрешкова, Фельберга и другие [22].

4.5. Другие методы, основанные на разделении конечномерной и бесконечномерной фазовых составляющих

Изложенный подход к конструированию численных методов решения ФДУ, опирающийся на разделение конечномерной и бесконечномерной составляющих в фазовой структуре системы (3.1), позволяет создавать (в сочетании с интерполяцией предыстории модели) полные аналоги всех известных для ОДУ численных алгоритмов.

Отметим следующие основные виды численных методов:

- 1) одношаговые k -этапные (методы типа Рунге–Кутты),
- 2) многошаговые,
- 3) использующие вычисление старших производных,
- 4) экстраполяционные,
- 5) циклические,
- 6) блочные,
- 7) гибридные,
- 8) многошаговые многоэтапные с использованием старших производных,
- 9) Нордсика,
- 10) методы решения уравнений второго порядка,
- 11) симплектические методы.

Заметим, что все они являются *позиционными*, т.е. в момент t можно определить часть траектории до этого момента, в отличие, например, от методов решения ФДУ путем сведения к интегральному уравнению, где решение определяется целиком. Поэтому перечисленные методы можно использовать в задачах позиционного управления системами ФДУ. Некоторые из них уже были подробно рассмотрены нами или упомянуты, в остальных обратим внимание лишь на идейную сторону.

1) Экстраполяционные методы. В разделе 3.3.3 будет изложено правило Рунге оценки погрешности и экстраполяция по Ричардсону при уменьшении шага в два раза. Этот процесс можно продолжать, уменьшая шаг в 2, 4, 8, 16,... раз. Последовательность Булирша 1, 2, 3, 4, 6, 8, 12, 16, 24,... является более экономичной [22]. Экстраполяционные методы имеют то преимущество, что у них на каждом шаге можно менять не только длину шага, но и порядок метода, благодаря чему можно

достигнуть высокой точности. Теоретической основой методов является асимптотическое представление глобальной погрешности, для ФДУ этот результат изложен в разделе 3.2.5.

2) Циклические методы. Несколько отличных друг от друга многошаговых методов используются циклически [22]. На их основе можно получить сходящиеся методы порядка $2k - 1$ (или $2k$ при четных k). Таким способом можно разрушить барьер Далквиста (см. раздел 3.1.6).

3) Блочные методы. Методами типа Рунге–Кутты одновременно вычисляется “блок” приближений $u_{l+1}, u_{l+2}, \dots, u_{l+k}$ [25]. Преимущество блочных методов заключается в том, что для них требуется меньше вычислений значений f , чем для традиционных методов Рунге–Кутты.

4) Гибридные методы. В линейных многошаговых методах используется информация лишь в точках $t_l = t_0 + l\Delta$, $l = 0, 1, 2, \dots$, тогда как в методах Рунге–Кутты может использоваться информация в других точках, например в $t_l + \Delta/2$. В гибридных методах объединяют свойства многошаговых методов и методов Рунге–Кутты, например k -шаговая гибридная формула

$$\sum_{i=0}^k \alpha_i u_{l+i} = \Delta \sum_{i=0}^k \beta_i f_{l+i} + \Delta \beta_\nu f_{l+\nu},$$

где ν – не целое. За счет выбора параметров ν и β_ν удается получить методы с лучшими (по сравнению с многошаговыми и Рунге–Кутты) свойствами [25], в частности, удалось разрушить барьер Бутчера (см. раздел 3.1.5).

5) Многошаговые многоэтапные методы с использованием старших производных. Эти методы представляют, видимо, наиболее широкий класс задаваемых в явном виде линейных методов. Используя обозначения раздела 3.1.9, их можно записать следующим образом:

$$u_{l+1}^i = \sum_{j=1}^k a_{ij} u_l^j + \sum_{r=1}^q \frac{\Delta^r}{r!} \sum_{j=1}^s b_{ij} j^r D^r x(t_l + c_j \Delta, v_l^j, v_{t_l + c_j \Delta}^j(\cdot)), \quad i = 1, \dots, k,$$

$$v_l^i = \sum_{j=1}^k \hat{a}_{ij} u_l^j + \sum_{r=1}^q \frac{\Delta^r}{r!} \sum_{j=1}^s \hat{b}_{ij} j^r D^r x(t_l + c_j \Delta, v_l^j, v_{t_l + c_j \Delta}^j(\cdot)), \quad i = 1, \dots, s.$$

Этапы $(u_l^i, i = 1, \dots, k)$ называются *внешними* и содержат всю необходимую информацию с предыдущего этапа, этапы $(v_l^i, i = 1, \dots, s)$ — *внутренними*. Предполагается, что выбран оператор интерполяции–экстраполяции предыстории модели $IE : \{u_i\}_l \rightarrow v(\cdot) \in Q[t_l - \tau, t_l + c\Delta]$, где $c = \max_i \{c_i\}$.

6) Методы решения уравнений второго порядка. В задачах механики чаще всего встречаются различные виды ОДУ второго порядка

$$\ddot{x} = f(t, x, \dot{x}).$$

Созданные для их решения численные методы — одни из самых старых, но в то же время они не потеряли своей актуальности, так как хорошо учитывают специфику объекта. Методы можно разделить на два основных класса:

- одношаговые многоэтапные (например, метод Нюстрема [22]),
- многошаговые (например, метод Штернера.

Аналогичные методы можно применять для численного решения ФДУ вида

$$\ddot{x}(t) = f(t, x(t), \dot{x}(t), x_t(\cdot), \dot{x}_t(\cdot)).$$

7) Симплектические или лиувиллевы методы. Методы, сохраняющие фазовый объем, называются лиувиллевыми. В частности, этими методами можно считать замкнутые траектории (циклы). Примером может служить пара методов: явный и неявный методы Эйлера, используемых по очереди. Близки к симплектическим *проекционные* методы [51].

Все эти и другие методы вкладываются в общую схему, изложенную в следующей главе.

Глава 5.

Общая схема численных методов

5.1. Введение

Рассмотрим начальную задачу для функционально-дифференциальных уравнений

$$\dot{x}(t) = f(t, x_t(\cdot)), \quad t \in [\alpha, \beta], \quad (5.1)$$

$$x(t) = \phi(t), \quad t \in [\alpha - \tau, \alpha]. \quad (5.2)$$

Здесь $\tau \geq 0$ — величина запаздывания, $x \in \mathbb{R}^n$ — фазовый вектор, \mathbb{R}^n — n -мерное евклидово пространство с нормой $\|\cdot\|$, $x_t(\cdot) = \{x(t+s), \alpha - \tau \leq t+s \leq t\}$ — предыстория фазового вектора к моменту t .

Пусть $X[t_1, t_2]$ — линейное нормированное пространство функций $x(t)$ со значениями в \mathbb{R}^n , определенных на $[t_1, t_2]$ и с нормой $\|x(\cdot)\|_{X[t_1, t_2]}$.

Будем предполагать, что $f(t, x_t(\cdot))$, отображающая $t \in [\alpha, \beta]$ и $x_t(\cdot) \in X[\alpha - \tau, t]$ в \mathbb{R}^n , и функция $\phi \in X[\alpha - \tau, \alpha]$ таковы, что задача (5.1) - (5.2) имеет единственное решение $x(\cdot) \in X[\alpha - \tau, \beta]$.

В частности, если $X[t_1, t_2] = C[t_1, t_2]$ — пространство непрерывных n -мерных функций с нормой $\|x(\cdot)\|_{C[t_1, t_2]} = \max_{t_1 \leq t \leq t_2} \|x(t)\|$, то существование и единственность решения задачи (5.1) - (5.2) гарантируются [24, с.55] следующими условиями:

A1: $f(t, x(\cdot))$ непрерывна по t при фиксированном $x(\cdot) \in C[\alpha - \tau, \beta]$

A2: $f(t, x(\cdot))$ липшицева по $x(\cdot)$, т.е. найдется L , такое, что

$\|f(t, x^1(\cdot)) - f(t, x^2(\cdot))\| \leq L \|x^1(\cdot) - x^2(\cdot)\|_{C[\alpha - \tau, t]}$ для всех $t \in [\alpha, \beta]$ и $x^1(\cdot), x^2(\cdot) \in C[\alpha - \tau, \beta]$.

Следующий случай, $X[t_1, t_2] = \mathbb{R}^n \times Q[t_1, t_2]$, где $Q[t_1, t_2]$ — пространство кусочно-непрерывных на $[t_1, t_2]$ n -мерных функций $x(\cdot)$ (непрерывных справа в точках разрыва) с нормой $\|x(\cdot)\|_{Q[t_1, t_2]} = \sup_{t_1 \leq t < t_2} \|x(t)\|$.

Введение этого пространства позволяет разделить конечномерную и бесконечномерную составляющие в предыстории фазового вектора и рассмотреть частный случай системы (5.1) - (5.2):

$$\dot{x} = f(t, x, y), \quad t \in [\alpha, \beta], \quad (5.3)$$

$$x(\alpha) = x^0, \quad y(\alpha) = y^0. \quad (5.4)$$

Здесь $x = x(t) \in \mathbb{R}^n$, $y = y(t) = \{x_t(s) = x(t+s), \alpha - \tau \leq t+s < t\} \in Q[\alpha - \tau, t]$.

Существование и единственность решения $x(\cdot) \in Q[\alpha - \tau, \beta]$ задачи (5.3) - (5.4) гарантируются (см. теорему ??) условиями 3.1 и 3.2.

Численные методы для задачи (5.1) - (5.2), особенно для систем с сосредоточенным запаздыванием, изучались во многих работах, см. обзоры [25, 30, 32, 37]. Специфика системы (5.3) - (5.4) позволила разработать одношаговые методы типа Рунге-Кутты и многошаговые методы, аналогичные соответствующим для обыкновенных дифференциальных уравнений.

Для обыкновенных дифференциальных уравнений имеется ряд работ, в которых методы типа Рунге-Кутты, многошаговые и другие объединяются в общие схемы.

Наиболее общей, видимо, следует считать схему [53], см. также [22], в которую вкладываются все известные автору методы. В данной главе эта схема [53] модифицируется для случая системы ФДУ. Главным моментом в модификации является введение промежуточного пространства между дискретной схемой и непрерывной функциональной системой, этот элемент назван интерполяцией. Кроме того, традиционное условие липшицевости функции, определяющей продвижение дискретной модели на шаг, заменяется на более слабое условие, названное квазиплипшицевостью.

5.2. Дискретная модель и порядок сходимости

Определение 5.1. Шагом сетки назовем число $\Delta > 0$, такое, что $\tau/\Delta = m$ — целое, $\{\Delta\}$ — множество шагов.

Определение 5.2. Пусть $t_0 \in [\alpha, \beta]$. Сеткой (равномерной) назовем конечный набор чисел $\Sigma_\Delta = \{t_i = t_0 + i\Delta \in [\alpha - \tau, \beta], i = -m, \dots, N\}$. Обозначим $\Sigma_\Delta^- = \{t_i \in \Sigma_\Delta, i \leq 0\}$, $\Sigma_\Delta^+ = \{t_i \in \Sigma_\Delta, i \geq 0\}$.

Определение 5.3. Дискретной моделью назовем всякую сеточную функцию $U: t_i \in \Sigma_\Delta \rightarrow u(t_i) = u_i \in U_i, i = -m, \dots, N$, где $U_i, i = -m, \dots, N$ — линейные нормированные пространства с нормами $\|\cdot\|_{U_i}$. В частности, во всех методах решения ОДУ и в большинстве методов решения ФДУ (см. ниже) $U_i = \mathbb{R}^q, i = -m, \dots, N$.

Определение 5.4. Для $l \geq 0$ предысторией дискретной модели к моменту t_l назовем множество $\{u_i\}_{-m}^l = \{u_i \in U_i, i = -m, \dots, l\}$.

Пусть V_l — линейные нормированные пространства с нормой $\|\cdot\|_V$.

Определение 5.5. Оператором интерполяции дискретной предыстории модели назовем функцию $I: I(\{u_i\}_{-m}^l) = v \in V_l$.

Определение 5.6. Будем говорить, что оператор интерполяции удовлетворяет условию квазиплипшицевости, если найдутся такие K_1 и $\omega_1(\Delta)$, что для всех $\Delta \geq 0$ и для всяких предысторий дискретной модели $\{u_i^1\}_{-m}^l$ и $\{u_i^2\}_{-m}^l$ выполняется неравенство:

$$\|v^1 - v^2\|_V \leq K_1 \max_{-m \leq i \leq l} \|u_i^1 - u_i^2\|_{U_i} + \omega_1(\Delta), \quad (5.5)$$

где $v^j = I(\{u_i^j\}_{-m}^l), j = 1, 2, \omega_1(\Delta) \rightarrow 0$ при $\Delta \rightarrow 0$.

Определение 5.7. Будем говорить, что оператор интерполяции удовлетворяет условию квазиплипшицевости с порядком p , если существует константа C такая, что

$$\omega_1(\Delta) \leq C\Delta^p.$$

Определение 5.8. Функцией точных значений назовем отображение Z :

$$Z(t_i, \Delta) = z_i \in U_i, i = -m, \dots, N. \quad (5.6)$$

Будем предполагать, что задание функции точных значений является следствием задания точного решения задачи (5.1) – (5.2), т.е. существует отображение $\Lambda : \mathbb{R}^n \rightarrow U_i$:

$$z_i = \Lambda x_i, \quad i = -m, \dots, N.$$

Для восстановления точного решения в узлах сетки нужно обратить оператор Λ ; в частности, если $U_i = \mathbb{R}^q$ и отображение Λ линейное, то $x_i = \Lambda^+ z_i$, $i = -m, \dots, N$, где Λ^+ - псевдообратная матрица отображения.

Определение 5.9. *Стартовыми значениями модели назовем функцию $U^0 : \Sigma_{\Delta}^- \times \{\Delta\} \rightarrow U_i$:*

$$u_i = U^0(t_i, \Delta), \quad i = -m, \dots, 0. \quad (5.7)$$

Определение 5.10. *Будем говорить, что стартовые значения модели имеют порядок p , если найдется константа C такая, что*

$$\| z_i - U^0(t_i, \Delta) \|_{U_i} \leq C \Delta^p \text{ для всех } i = -m, \dots, 0.$$

Определение 5.11. *Формулой продвижения модели на шаг назовем алгоритм*

$$u_{l+1} = S_l u_l + \Delta \Phi(t_l, I(\{u_i\}_{-m}^l), \Delta), \quad (5.8)$$

где $\Phi : \Sigma_{\Delta}^+ \times V_l \times \{\Delta\} \rightarrow U_{l+1}$, $S_l : U_l \rightarrow U_{l+1}$ – линейный оператор.

Таким образом, дискретная модель (синоним численный метод или просто метод) определяется функцией точных значений (5.6), стартовыми значениями (5.7), формулой продвижения на шаг (5.8) и оператором интерполяции.

Будем предполагать, что функция $\Phi(t_l, v, \Delta)$ в (5.8) квазилипшицева по второму аргументу, т.е. найдутся такие K_2 и $\omega_2(\Delta)$, что для всех $t_l \in \Sigma_{\Delta}^+$, $\Delta \in \{\Delta\}$, $v^1, v^2 \in V_l$ выполняется

$$\| \Phi(t_l, v^1, \Delta) - \Phi(t_l, v^2, \Delta) \|_{U_{l+1}} \leq K_2 \| v^1 - v^2 \|_{V_l} + \omega_2(\Delta), \quad (5.9)$$

где $\omega_2(\Delta) \rightarrow 0$ при $\Delta \rightarrow 0$.

Будем говорить, что функция Φ удовлетворяет условию квазилипшицевости с порядком p , если существует константа C такая, что

$$\omega_2(\Delta) \leq C \Delta^p.$$

Замечание 5.1. Из квазилипшицевости I и Φ следует квазилипшицевость суперпозиции $\Phi(t_l, I(\{u_i\}_{-m}^l), \Delta)$ по дискретной предыстории модели. В самом деле,

$$\begin{aligned} & \| \Phi(t_l, I(\{u_i^1\}_{-m}^l), \Delta) - \Phi(t_l, I(\{u_i^2\}_{-m}^l), \Delta) \|_{U_{l+1}} \leq \\ & \leq K_2 \| I(\{u_i^1\}_{-m}^l) - I(\{u_i^2\}_{-m}^l) \|_{V_l} + \omega_2(\Delta) \leq \\ & \leq K_2 K_1 \max_{-m \leq i \leq l} \| u_i^1 - u_i^2 \|_{U_i} + K_2 \omega_1(\Delta) + \omega_2(\Delta), \end{aligned}$$

откуда, обозначив $K = K_2 K_1$ и $\omega(\Delta) = K_2 \omega_1(\Delta) + \omega_2(\Delta)$, получим требуемое.

Определение 5.12. Будем говорить, что метод сходится, если для погрешности метода $\delta_l = z_l - u_l$ выполняется

$$\max_{-m \leq l \leq N} \|\delta_l\| \rightarrow 0 \text{ при } \Delta \rightarrow 0.$$

Определение 5.13. Будем говорить, что метод сходится с порядком p , если существует константа C такая, что

$$\|\delta_l\| \leq C\Delta^p \text{ для всех } l = -m, \dots, N.$$

Индексы у норм в дальнейшем будем опускать.

Определение 5.14. Назовем метод (5.8) устойчивым, если произведение операторов S_i , $i = l, l-1, \dots, k$ равномерно ограничено для всех $l \geq 0$, $0 \leq k \leq l$, т.е. существует \tilde{S} такое, что

$$\|S^{l,k}\| \leq \tilde{S}, \quad S^{l,k} = S_l \cdot S_{l-1} \cdots S_k.$$

Определение 5.15. Погрешностью аппроксимации (невязкой) метода назовем сеточную функцию $\Sigma_\Delta \rightarrow U_l$:

$$\begin{aligned} d_l &= z_l - U^0(t_l, \Delta), \quad l = -m, \dots, 0, \\ d_{l+1} &= z_{l+1} - S_l z_l - \Delta \Phi(t_l, I(\{z_i\}_{-m}^l), \Delta), \quad l = 0, \dots, N-1. \end{aligned} \quad (5.10)$$

Определение 5.16. Будем говорить, что метод (5.8) имеет порядок погрешности аппроксимации p , если существует константа C такая, что

$$\|d_l\| \leq C\Delta^{p+1} \text{ для всех } l = 1, \dots, N.$$

Справедлива следующая основная теорема.

Теорема 5.1. Пусть метод (5.8) устойчив, функция Φ удовлетворяет условию квазилипицевости с порядком $p_1 > 0$, оператор интерполирования I удовлетворяет условию квазилипицевости с порядком $p_2 > 0$, стартовые значения имеют порядок $p_3 > 0$, погрешность аппроксимации имеет порядок $p_4 > 0$, тогда метод сходится, причем порядок сходимости не меньше минимума из p_1, p_2, p_3, p_4 .

Доказательство опирается на следующую лемму.

Лемма 5.1. Пусть метод (5.8) устойчив, функция Φ удовлетворяет условию квазилипицевости (5.9), оператор интерполирования I удовлетворяет условию квазилипицевости (5.5) и пусть \hat{u}_l , $l = -m, \dots, N$, — решение возмущенного уравнения

$$\hat{u}_l = U^0(t_l, \Delta) + r_l, \quad l = -m, \dots, 0,$$

$$\hat{u}_{l+1} = S_l \hat{u}_l + \Delta \Phi(t_l, I(\{\hat{u}_i\}_{-m}^l), \Delta) + r_{l+1}, \quad l = 0, \dots, N-1, \quad (5.11)$$

где $r_l \in U_l$, $l = -m, \dots, N$, — возмущения, тогда имеют место оценки

$$\max_{0 \leq l \leq N} \|\hat{u}_l - u_l\| \leq A e^{K\tilde{S}(\beta-\alpha)}, \quad (5.12)$$

$$\|R\|_S \leq A_1 \max_{0 \leq l \leq N} \|\hat{u}_l - u_l\| + A_2, \quad (5.13)$$

где

$$\begin{aligned} A &= R_0 + \tilde{S}(\beta - \alpha)\omega(\Delta) + \|R\|_S, \\ A_1 &= 1 + K\tilde{S}(\beta - \alpha), \\ A_2 &= (1 + K\tilde{S}(\beta - \alpha))R_0 + \tilde{S}(\beta - \alpha)\omega(\Delta), \\ R_0 &= \max\{\|r_i\|, -m \leq i \leq 0\}, \\ \|R\|_S &= \max_{0 \leq l \leq N} \left\| \sum_{j=0}^l S^{l,j} r_j \right\|, \end{aligned}$$

K и $\omega(\Delta)$ определены в замечании 5.1, обозначения $S^{l,j}$ и \tilde{S} взяты из определения устойчивости.

Доказательство леммы.

Обозначим $\delta_l = \hat{u}_l - u_l$, $l = -m, \dots, N$, тогда для $l = 0, \dots, N-1$ имеем

$$\delta_{l+1} = \sum_{j=0}^l S^{l,j} \Delta \hat{\delta}_j + \sum_{j=0}^{l+1} S^{l+1,j} r_j, \quad (5.14)$$

где

$$\hat{\delta}_l = \Phi(t_l, I(\{\hat{u}_i\}_{-m}^l), \Delta) - \Phi(t_l, I(\{u_i\}_{-m}^l), \Delta).$$

Из замечания 5.1 следует

$$\|\hat{\delta}_l\| \leq K \max_{-m \leq i \leq l} \{\|\delta_i\|\} + \omega(\Delta), \quad (5.15)$$

тогда из определения устойчивости получим

$$\|\delta_{l+1}\| \leq \tilde{S}K\Delta \sum_{j=0}^l \max_{-m \leq i \leq j} \{\|\delta_i\|\} + \tilde{S}\Delta \sum_{j=0}^l \omega(\Delta) + \left\| \sum_{j=0}^{l+1} S^{l+1,j} r_j \right\|$$

и

$$\|\delta_{l+1}\| \leq \tilde{S}K\Delta \sum_{j=0}^l \max_{-m \leq i \leq j} \{\|\delta_i\|\} + \tilde{S}(\beta - \alpha)\omega(\Delta) + \|R\|_S. \quad (5.16)$$

Индукцией по $l = 1, \dots, N$ докажем оценку

$$\begin{aligned} \|\delta_l\| &\leq (1 + \tilde{S}K\Delta)^l (R_0 + \tilde{S}(\beta - \alpha)\omega(\Delta) + \|R\|_S) = \\ &= A(1 + \tilde{S}K\Delta)^l. \end{aligned} \quad (5.17)$$

База индукции. Из (5.8) и (5.11) следует

$$\|\delta_1\| \leq \tilde{S}K\Delta R_0 + \tilde{S}(\beta - \alpha)\omega(\Delta) + \|R\|_S \leq$$

$$\leq (1 + \tilde{S}K\Delta)(R_0 + \tilde{S}(\beta - \alpha)\omega(\Delta) + \|R\|_S).$$

Шаг индукции. Пусть оценка (5.17) верна для всех индексов от 1 до l . Покажем, что оценка справедлива и для $l+1$. Зафиксируем $j \leq l$. Пусть i_0 — индекс, на котором достигается

$$\max_{-m \leq i \leq j} \{\|\delta_i\|\}.$$

Возможны две ситуации:

а) $i_0 \leq 0$, тогда

$$\max_{-m \leq i \leq j} \{\|\delta_i\|\} = \|\delta_{i_0}\| \leq R_0 \leq A(1 + \tilde{S}K\Delta)^j;$$

б) $1 \leq i_0 \leq j$, тогда по индуктивному предположению

$$\max_{-m \leq i \leq j} \{\|\delta_i\|\} = \|\delta_{i_0}\| \leq A(1 + \tilde{S}K\Delta)^{i_0} \leq A(1 + \tilde{S}K\Delta)^j.$$

Из полученной оценки и (5.16) вытекает

$$\|\delta_{l+1}\| \leq \tilde{S}K\Delta \sum_{j=0}^l A(1 + \tilde{S}K\Delta)^j + A = A(1 + \tilde{S}K\Delta)^{l+1}.$$

Таким образом, оценка (5.17) доказана и из нее получаем утверждение (5.12) леммы:

$$\|\delta_l\| \leq A(1 + \tilde{S}K\Delta)^l \leq Ae^{K\tilde{S}(\beta-\alpha)}.$$

Из (5.14) следует, что при всех $0 \leq l \leq N-1$ выполняется

$$\left\| \sum_{j=0}^{l+1} S^{l+1,j} r_j \right\| \leq \|\delta_{l+1}\| + \tilde{S}\Delta l \max_{0 \leq i \leq l} \|\hat{\delta}_i\|,$$

тогда из (5.15) имеем

$$\left\| \sum_{j=0}^{l+1} S^{l+1,j} r_j \right\| \leq (1 + \tilde{S}K\Delta l) \max_{-m \leq i \leq l} \|\delta_i\| + \tilde{S}\Delta l \omega(\Delta),$$

или

$$\|R\|_S \leq (1 + \tilde{S}K(\beta - \alpha)) \left(\max_{-m \leq i \leq N} \|\delta_i\| + R_0 \right) + \tilde{S}(\beta - \alpha)\omega(\Delta),$$

что доказывает оценку (5.13) □

Доказательство теоремы 5.1.

Возьмем в (5.11) возмущения, равные погрешности аппроксимации, т.е. $r_l = d_l$, тогда из (5.10) возмущенная модель \hat{u}_l определяется

$$\text{для } l = -m, \dots, 0 : \hat{u}_l = U^0(t_l, \Delta) + z_l - U^0(t_l, \Delta) = z_l,$$

$$\text{для } l = 0, \dots, N-1 : \hat{u}_{l+1} = S_l \hat{u}_l + \Delta \Phi(t_l, I(\{\hat{u}_i\}_{-m}^l), \Delta) - S_l z_l - \Delta \Phi(t_l, I(\{z_i\}_{-m}^l), \Delta), \text{ и в этом случае индуктивно получаем } \hat{u}_{l+1} = z_{l+1}, l = 0, \dots, N-1.$$

Тогда из утверждения (5.12) леммы 5.1 следует

$$\|z_l - u_l\| \leq (R_0 + \tilde{S}(\beta - \alpha)\omega(\Delta) + \|R\|_S) e^{K\tilde{S}(\beta-\alpha)}$$

при $l = 1, \dots, N$, откуда, с учетом того, что

$$\|R\|_S \leq \tilde{S}(\beta - \alpha) \max_{0 \leq i \leq N} \|r_i\| / \Delta$$

вытекает заключение теоремы. \square

5.3. Методика классификации численных моделей ФДУ

Как показывает общее описание модели, численные методы отличаются:

- а) по выбору сетки и пространств дискретных моделей;
- б) по способу интерполяции предыстории дискретной модели;
- в) по выбору стартовых значений модели;
- г) по выбору формулы продвижения на шаг.

Помимо этого для введения понятия погрешности необходимо описание функции точных значений. Основным моментом в этой классификации, как и для ОДУ, является последний пункт — формула продвижения на шаг, которую часто и называют методом. Все известные как для ОДУ, так и для ФДУ, методы вкладываются в эту схему. Рассмотрим некоторые примеры вложений.

Явные методы Рунге–Кутты. Явные k -этапные методы типа Рунге–Кутты, изученные для ФДУ в предыдущей главе, существенно опираются на разделение в объекте конечномерной и бесконечномерной составляющих, т.е. рассматриваются системы (5.3) – (5.4) с условиями 3.1 и 3.2.

Опишем модель, вкладывая ее в общую схему. Пусть дискретная модель конечномерна $U_l = \mathbb{R}^q$, причем ее размерность совпадает с размерностью фазового вектора, т.е. $q = n$. Сетки будем рассматривать только такие, для которых $t_0 = \alpha$. Пусть $\Delta > 0$, $m = \tau / \Delta$.

Функция точных значений пусть задается фазовым вектором $Z(t_i, \Delta) = z_i = x(t_i)$, $i = -m, \dots, N$, т.е. отображение Λ — тождественный оператор.

Стартовые значения модели пусть совпадают с точными значениями $u_l = x(t_l)$, $l = -m, \dots, 0$. Таким образом, начальная погрешность аппроксимации $d_l = 0$, $l = -m, \dots, 0$.

Специфика методов типа Рунге–Кутты состоит в том, что для корректного определения формулы продвижения на шаг в момент t_l необходимо произвести не только интерполяцию дискретной предыстории с сетки на отрезок $[t_l - \tau, t_l]$, но и произвести экстраполяцию предыстории на отрезок $[t_l, t_l + a\Delta]$, $a > 0$. Пусть $a > 0$ зафиксировано, возьмем в качестве интерполяционного пространства $V = Q[t_l - \tau, t_l + a\Delta]$. В разделе 3.1.2 указаны способы построения оператора интерполяции (в случае методов Рунге–Кутта он состоит из собственно интерполяции и из экстраполяции). Например простейший из них — кусочно-постоянная интерполяция и экстраполяция: $I(\{u_i\}_{-m}^l) = I(\{u_i\}_{l-m}^l) = v(\cdot) \in Q[t_l - \tau, t_l + a\Delta]$,

$$v(t) = \begin{cases} u_i, & t \in [t_i, t_{i+1}), \quad i = l - m, \dots, l - 1, \\ u_l, & t \in [t_l, t_l + a\Delta). \end{cases}$$

Этот оператор липшицев с константой $K_1 = 1$.

Будем в дальнейшем предполагать, что оператор I согласован, т.е. $v(t_i) = u_i$ для всех $i = l - m, \dots, l$, $v(\cdot) = I(\{u_i\}_{-m}^l)$

Формула продвижения на шаг задается соотношением

$$u_{l+1} = u_l + \Delta \sum_{i=1}^k \sigma_i \Phi_i(t_l, I(\{u_i\}_{-m}^l), \Delta), \quad (5.18)$$

где

$$\Phi_1(t_l, I(\{u_i\}_{-m}^l), \Delta) = f(t_l, u_l, v_{t_l}(\cdot)), \quad (5.19)$$

$$\begin{aligned} \Phi_i(t_l, I(\{u_i\}_{-m}^l), \Delta) = \\ = f(t_l + a_i \Delta, u_l + \Delta \sum_{j=1}^{i-1} b_{ij} \Phi_j(t_l, I(\{u_i\}_{-m}^l), \Delta), v_{t_l + a_i \Delta}(\cdot)), \end{aligned} \quad (5.20)$$

σ_i, a_i, b_{ij} ($i = 1, \dots, k, j = 1, \dots, i-1$) — параметры метода, $|a_i| \leq a, i = 1, \dots, k$.

Лемма 5.2. Функция $\Phi(t_l, I(\{u_i\}_{-m}^l), \Delta) = \sum_{i=1}^k \sigma_i \Phi_i(t_l, I(\{u_i\}_{-m}^l), \Delta)$ липшицева по второму аргументу.

Доказательство леммы.

Пусть $v^1 = I(\{u_i^1\}_{-m}^l), v^2 = I(\{u_i^2\}_{-m}^l)$. Докажем индукцией по i липшицевость с константой L_i функций $\Phi_i(t_l, v, \Delta)$ по второму аргументу.

База индукции. Так как функция f удовлетворяет условию Липшица, а оператор I согласован, то

$$\begin{aligned} \|\Phi_1(t_l, v^2, \Delta) - \Phi_1(t_l, v^1, \Delta)\| &= \|f(t_l, u_l^2, v_{t_l}^2(\cdot)) - f(t_l, u_l^1, v_{t_l}^1(\cdot))\| \leq \\ &\leq L\|u_l^2 - u_l^1\| + M\|v^2 - v^1\|_{V_l} \leq (L + M)\|v^2 - v^1\|_{V_l}, \end{aligned}$$

т.е. $L_1 = L + M$.

Шаг индукции. В силу условия Липшица и индуктивного предположения имеем

$$\begin{aligned} \|\Phi_i(t_l, v^2, \Delta) - \Phi_i(t_l, v^1, \Delta)\| &= \|f(t_l + a_i \Delta, u_l^2 + \Delta \sum_{j=1}^{i-1} b_{ij} \Phi_j(t_l, v^2(\cdot), \Delta), \\ &\quad v_{t_l + a_i \Delta}^2(\cdot)) - f(t_l + a_i \Delta, u_l^1 + \Delta \sum_{j=1}^{i-1} b_{ij} \Phi_j(t_l, v^1(\cdot), \Delta), \\ &\quad v_{t_l + a_i \Delta}^1(\cdot))\| \leq \\ &\leq L\Delta \sum_{j=1}^{i-1} |b_{ij}| \|\Phi_j(t_l, v^2(\cdot), \Delta) - \Phi_j(t_l, v^1(\cdot), \Delta)\| + M\|v^2 - v^1\|_{V_l} \leq \\ &\leq L\Delta \sum_{j=1}^{i-1} |b_{ij}| L_j \|v^2 - v^1\|_{V_l} + M\|v^2 - v^1\|_{V_l}. \end{aligned}$$

Таким образом,

$$L_i = (L\Delta \sum_{j=1}^{i-1} |b_{ij}| L_j + M).$$

Из липшицевости функций $\Phi_i(t_l, v, \Delta)$ следует липшицевость функции $\Phi(t_l, v, \Delta)$ по второму аргументу с константой $L_\Phi = \sum_{i=1}^k |\sigma_i| L_i$. \square

Условия устойчивости в ЯРК-методах выполняются, так как S_l — тождественный оператор.

Таким образом, из теоремы 5.1 вытекает, что метод Рунге–Кутты имеет порядок сходимости p , если локальный порядок его погрешности аппроксимации равен $p + 1$.

Локальный порядок погрешности аппроксимации (невязки) в общей схеме определяется через функцию d_l в (5.10), т.е. зависит от формулы продвижения на шаг и от способа интерполяции.

В разделе 3.1.3, где исследован явный метод Рунге–Кутты, в силу того, что размерности модели и фазового вектора совпадают, невязка определяется иначе, чем в (5.10):

Определение 5.17. *Невязкой (без интерполяции) ЯРК-метода назовем*

$$\tilde{d}_{l+1} = x_{l+1} - x_l - \Delta \Phi(t_l, x(\cdot), \Delta), \quad l = 0, \dots, N-1,$$

где

$$\begin{aligned} \Phi(t_l, x(\cdot), \Delta) &= \sum_{i=1}^k \sigma_i \Phi_i(t_l, x(\cdot), \Delta), \\ \Phi_1(t_l, x(\cdot), \Delta) &= f(t_l, x_l, x_{t_l}(\cdot)), \\ \Phi_i(t_l, x(\cdot), \Delta) &= f(t_l + a_i \Delta, x_l + \Delta \sum_{j=1}^{i-1} b_{ij} \Phi_j(t_l, x(\cdot), \Delta), x_{t_l + a_i \Delta}(\cdot)). \end{aligned}$$

Будем говорить, что ЯРК-метод имеет *локальный порядок невязки (без интерполяции)* $p + 1$, если существует константа C , такая, что

$$\|\tilde{d}_l\| \leq C \Delta^{p+1} \quad \text{для всех } l = 1, \dots, N. \quad (5.21)$$

Локальный порядок невязки (без интерполяции) определяется набором параметров $\sigma_i, a_i, b_{ij} (i = 1, k(1), j = 1, i-1(1))$, причем по тем же правилам, что и для ОДУ, это проверяется с помощью техники i -гладкого анализа.

Определение 5.18. *Оператор интерполяции I имеет порядок p на классе функций X , содержащих точное решение, если найдется константа C , такая, что для всех $x(\cdot) \in X$, $l = 0, \dots, N$ выполняется*

$$\|I\{x_i\}_{l-m}^l - x(\cdot)\|_{V_l} \leq C \Delta^p. \quad (5.22)$$

Например, оператор интерполирования вырожденными сплайнами $p - 1$ степени имеет порядок интерполяции p на классе функций, непрерывно дифференцируемых p раз.

Лемма 5.3. Если ЯРК-метод имеет невязку (без интерполяции) порядка $p + 1$, и оператор интерполяции имеет порядок p на классе функций, содержащих точное решение, то метод имеет локальный порядок аппроксимации $p + 1$ в смысле определения (5.22).

Доказательство леммы.

Согласно определению невязки (5.10), для явных методов Рунге–Кутты имеем для всех $l = 0, \dots, N - 1$

$$\begin{aligned} \|d_{l+1}\| &= \|x_{l+1} - x_l - \Delta\Phi(t_l, I(\{x_i\}_{-m}^l), \Delta)\| \leq \\ &\leq \|x_{l+1} - x_l - \Delta\Phi(t_l, x(\cdot), \Delta)\| + \|\Delta\Phi(t_l, I(\{x_i\}_{-m}^l), \Delta) - \Delta\Phi(t_l, x(\cdot), \Delta)\|, \end{aligned}$$

тогда из (5.21), леммы 5.2 и (5.22) имеем

$$\|d_{l+1}\| \leq C\Delta^{p+1}.$$

□

Непрерывные одношаговые методы. В работах [37, 54] для численного решения задачи (5.1) – (5.2) были сконструированы методы, которые дают приближения точного решения не только в точках разбиения, но и во всех промежуточных точках.

Пусть сетка $t_l = \alpha + l\Delta$, $l = 0, 1, \dots, N$, $\Delta = (\beta - \alpha)/N > 0$.

Определение 5.19. Непрерывным одношаговым методом назовем метод

$$u(t) = \phi(t), \quad t \in [\alpha - \tau, \alpha], \quad (5.23)$$

$$u(t_l + r\Delta) = u(t_l) + r\Delta\Psi(t_l, u_{t_l}(\cdot), r), \quad 0 \leq r \leq 1, \quad (5.24)$$

где $\Psi(t, u_t(\cdot), r)$ — отображение, ставящее в соответствие $t \in [\alpha, \beta]$, $u_t(\cdot) \in C[\alpha - \tau, t]$ и $r \in [0, 1]$ вектор из \mathbb{R}^n .

В работе [54] для доказательства сходимости метода (5.23) – (5.24) с порядком p требуется, чтобы отображение $\Psi(t, u_t(\cdot), r)$ было непрерывно по r , липшицево по $u_t(\cdot)$, и обеспечивало порядок невязки p :

$$\left\| \frac{x(t_l + r\Delta) - x(t_l)}{\Delta} - \Psi(t_l, x_{t_l}(\cdot), r) \right\| \leq C\Delta^p, \quad (5.25)$$

для всех $r \in [0, 1]$, $l = 0, 1, \dots, N - 1$.

Покажем, что методы (5.23) – (5.24) вкладываются в общую схему.

Введем пространства $U_l = C[\alpha - \tau, t_l]$ и модели — наборы $u_l = u_{t_l}(\cdot) \in U_l$. Интерполяционные пространства для таких моделей не нужны и положим $V_l = U_l$.

Определим оператор $S_l: U_l \rightarrow U_{l+1}$ следующим образом:

$$S_l u(\cdot) = \hat{u}(t) = \begin{cases} u(t), & t < t_l, \\ u(t_l), & t \geq t_l. \end{cases}$$

Оператор S_l — линейный и $\|S_l\| = 1$.

Функция точных значений задаётся точным решением задачи (5.1) – (5.2): $z_l = x_{t_l}(\cdot) \in U_l = C[\alpha - \tau, t_l]$.

Формула продвижения на шаг дается соотношением

$$u_{l+1} = S_l u_l + \Delta \Phi(t_l, u_l, \Delta), \quad (5.26)$$

где

$$\Phi(t_l, u_l, \Delta) = \tilde{u}(t) = \begin{cases} 0, & t < t_l, \\ (t - t_l) \Psi(t_l, u_{t_l}(\cdot), \frac{t-t_l}{\Delta}), & t_l \leq t \leq t_{l+1}. \end{cases}$$

В силу введенных u_l , S_l и Φ , формула продвижения модели на шаг представляет собой эквивалентную форму записи метода (5.24). Так как $\|S_l\| = 1$, то метод устойчив.

Если Ψ липшицева по второму аргументу, то Φ также липшицева по второму аргументу. Если Ψ имеет порядок аппроксимации p в смысле (5.25), то погрешность аппроксимации метода (5.26) также p , и из теоремы 5.1 следует сходимость с порядком p .

Непрерывные методы до четвертого порядка сходимости включительно построены в [54]. Эти методы требуют на одном шаге большего числа вычислений значений функционала правой части системы, чем методы типа Рунге–Кутты соответствующих порядков (взамен отсутствует процедура интерполяции). Так, в методе четвертого порядка требуется вычислять шесть значений функционала f .

Непрерывные методы четвертого и пятого порядков Дормана–Принса [22] послужили основой для конструирования алгоритмов с автоматическим выбором шага, применимых как для решения ОДУ, так и ФДУ.

Многошаговые и другие методы. Пусть $k \geq 1$ — целое число. Возьмем $t_0 = \alpha + (k-1)\Delta$, где $\Delta > 0$. Пусть в моменты $t_l = t_0 + l\Delta$ приближенное значение решения системы (5.3)–(5.4) обозначено через $u_l^1 = u^1(t_l) \in \mathbb{R}^n$. Пусть задан на множестве $\{u_i^1\}_{l-m}^l$, m $= \tau/\Delta$ некоторый оператор интерполяции $I(\{u_i^1\}_{l-m}^l)$ в пространство $Q[t_l - \tau, t_l]$.

Явным k -шаговым методом назовем (см. раздел 3.1.6) алгоритм

$$u_l^1 = \sum_{j=1}^k a_j u_{l-j}^1 + \Delta \sum_{j=1}^k b_j f(t_{l-j}, u_{l-j}^1, I(\{u_i^1\}_{l-j-m}^{l-j})). \quad (5.27)$$

Как и в случае ОДУ, многошаговые методы сводятся к одношаговым путем повышения размерности модели.

Введем вектор u_l из пространства $U_l = U = \mathbb{R}^q$, $q = kn$:

$$u_l = (u_l^1, u_{l-1}^1, \dots, u_{l-k+1}^1)^t = (u_l^1, u_l^2, \dots, u_l^k)^t,$$

и пусть $u(t) \in Q[t_l - \tau - k\Delta, t_l]$ — результат действия оператора интерполяции $I(\{u_i^1\}_{l-m}^l) = I(\{u_i^1\}_{l-m-k}^l)$.

Тогда метод (5.27) можно переписать в виде

$$u_{l+1} = S u_l + \Delta \Phi(t_l, I(\{u_i^1\}_{-m}^l), \Delta), \quad (5.28)$$

где

$$S = \begin{pmatrix} a_1 & a_2 & \dots & a_{k-1} & a_k \\ 1 & 0 & \dots & 0 & 0 \\ \cdot & \cdot & \dots & \cdot & \cdot \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix}, \quad \Phi(t_l, I(\{u_i^1\}_{-m}^l), \Delta) = \begin{pmatrix} \Phi^1 \\ 0 \\ \cdot \\ 0 \end{pmatrix},$$

$$\Phi^1 = \sum_{j=1}^k b_j f(t_{l-j+1}, u_l^j, u_{t_{l-j+1}}(\cdot)).$$

Собственные числа матрицы S являются корнями характеристического уравнения

$$\lambda^k + a_1 \lambda^{k-1} + \dots + a_k = 0,$$

поэтому метод устойчив тогда и только тогда, когда все корни этого уравнения по модулю не больше 1, а среди корней, равных по модулю 1, нет кратных.

Важно отметить, что для систем вида (5.3)–(5.4) указанный подход позволяет использовать все разнообразие численных методов, известных для ОДУ [22, 25]. К числу таких методов можно отнести явные и неявные одношаговые и многошаговые, использующие вычисление старших производных, экстраполяционные, циклические, блочные, гибридные, методы типа Нордсика, симплектические методы и различные варианты их комбинаций.

5.4. Необходимые и достаточные условия сходимости с порядком p

В этом разделе будем предполагать, что пространство дискретных моделей конечномерно: $U_l = \mathbb{R}^q$ для всех $l = -m, \dots, N$, и пусть в пошаговой формуле (5.8) линейные операторы S_l задаются матрицей $S \in \mathbb{R}^{q \times q}$, т.е.

$$u_{l+1} = Su_l + \Delta \Phi(t_l, I(\{u_i\}_{-m}^l), \Delta). \quad (5.29)$$

Во всех указанных выше методах, кроме непрерывных, эти условия выполняются.

Если невязка d_l стремится к нулю при Δ стремящемся к нулю, при всех $l = 1, \dots, N$, то выполняется равенство $Z(t, 0) = SZ(t, 0)$, и единица является собственным числом матрицы S . Будем считать это условие выполненным. Если, кроме того, метод устойчив, то матрица S не может иметь собственных чисел вне единичной окружности, а собственные числа равные по модулю единице не могут порождать жордановых серий. Обозначим собственные числа, равные по модулю единице через $\lambda_1 = 1, \lambda_2, \dots, \lambda_k$. Тогда жорданова каноническая форма S будет блочно-диагональной

$$S = T \text{diag} \left\{ \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}, \begin{pmatrix} \lambda_2 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & \lambda_2 \end{pmatrix}, \dots, \right. \\ \left. \begin{pmatrix} \lambda_k & 0 & \dots & 0 \\ 0 & \lambda_k & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & \lambda_k \end{pmatrix}, Y \right\} T^{-1},$$

T – невырожденная матрица. Таким образом,

$$S = E_1 + \lambda_2 E_2 + \dots + \lambda_k E_k + \tilde{E},$$

где

$$\begin{aligned} E_1 &= T \operatorname{diag}\{E, 0, 0, \dots, 0\} T^{-1}, \\ E_2 &= T \operatorname{diag}\{0, E, 0, \dots, 0\} T^{-1}, \\ &\dots \\ \tilde{E} &= T \operatorname{diag}\{0, 0, 0, \dots, Y\} T^{-1}, \end{aligned}$$

E – единичная матрица.

Определение 5.20. Метод (5.29) согласован с порядком p , если найдется C такое, что невязка удовлетворяет соотношениям

$$\begin{aligned} \|d_l\| &\leq C\Delta^p \text{ при } l = 0, \dots, N, \\ \|E_1(d_0 + d_1 + \dots + d_l) + d_{l+1}\| &\leq C\Delta^p \text{ при } l = 0, \dots, N-1. \end{aligned}$$

Если метод имеет порядок невязки p , то он согласован с порядком p . Обратное неверно, примеры таких методов см. в [53].

Определение 5.21. Будем говорить, что метод (5.29) сильно устойчив, если матрица S имеет все собственные числа по модулю строго меньше 1, за исключением простого корня $\lambda_1 = 1$.

Теорема 5.2. Пусть метод (5.29) сильно устойчив, функция Φ удовлетворяет условию квазилипицевости с порядком $p > 0$, оператор интерполирования I удовлетворяет условию квазилипицевости с порядком p , стартовые значения имеют порядок p . Тогда метод сходится с порядком p тогда и только тогда, когда он согласован с порядком p .

Доказательство. Пусть метод сходится с порядком p . В силу соотношения $E_1 S = E_1$ выполняется

$$E_1(d_0 + \dots + d_l) + d_{l+1} = \sum_{j=0}^{l+1} S^{l+1-j} d_j - (S - E_1) \sum_{j=0}^l S^{l-j} d_j,$$

откуда для $l \leq N-1$

$$\|E_1(d_0 + \dots + d_l) + d_{l+1}\| = (1 + \|S - E_1\|) \|D\|_S,$$

где $\|S\|$ – подчиненная норма матрицы,

$$\|D\|_S = \max_{0 \leq l \leq N} \left\| \sum_{j=0}^l S^{l-j} d_j \right\|.$$

Возьмем в (5.11) возмущения, равные невязке, т.е. $r_l = d_l$, тогда из оценки (5.13) леммы 3.1 получаем согласованность с порядком p .

Обратно, пусть метод согласован с порядком p . Так как метод сильно устойчив, то $S = E_1 + \tilde{E}$ и для $0 \leq j \leq l-1$ выполняется

$$S^{l-j} = E_1 + \tilde{E}^{l-j}. \quad (5.30)$$

Тогда при $r_l = d_l$

$$\begin{aligned} \|R\|_S &= \max_{0 \leq l \leq N} \left\| \sum_{j=0}^l S^{l-j} d_j \right\| = \max_{0 \leq l \leq N} \left\| \sum_{j=0}^{l-1} (E_1 d_j + \tilde{E}^{l-j} d_j) + d_l \right\| \leq \\ &\leq \max_{0 \leq l \leq N-1} \left\| \sum_{j=0}^l E_1 d_j + d_{l+1} \right\| + \frac{1}{1 - \|\tilde{E}\|} \max_{0 \leq l \leq N-1} \|d_l\|. \end{aligned}$$

В силу условия согласованности и того, что $\|\tilde{E}\| < 1$, выполняется $\|R\|_S = O(\Delta^p)$, что вместе с оценкой (5.12) леммы 3.1 показывает сходимость метода с порядком p . \square

Как и в случае ОДУ, условие сильной устойчивости метода в теореме 5.2 можно ослабить до условия устойчивости, если невязка обладает свойством гладкости.

Определение 5.22. Будем говорить, что невязка обладает свойством гладкости с порядком p , если невязку можно разложить в ряд Тейлора по степеням Δ :

$$d_{l+1} = d^0(t_l) + d^1(t_l)\Delta + \dots + d^p(t_l)\Delta^p + O(\Delta^p), \quad l = 0, \dots, N-1,$$

где $d^j(t)$ – непрерывно дифференцируемые $p-j+1$ раз функции.

Свойство гладкости невязки определяется гладкостью функции точных значений и функционала Φ в формуле продвижения модели на шаг, что в свою очередь определяется гладкостью функционала правой части системы ФДУ и её решения. В случае системы (5.3) - (5.4) эти условия могут быть выражены с помощью техники i -гладкого анализа.

Лемма 5.4. Если невязка обладает свойством гладкости с порядком p , то метод согласован с порядком p тогда и только тогда, когда

$$\|d_l\| \leq C\Delta^p \text{ при } l = 0, \dots, N, \quad \text{и } E_1 d^p(t) = 0 \text{ для } t \in (\alpha, \beta]. \quad (5.31)$$

Доказательство. Из условия гладкости невязки с порядком p и (5.31) следует

$$d_j = d^p(t_{j-1})\Delta^p + O(\Delta^p), \quad E_1 d_j = O(\Delta^p), \quad j = 1, \dots, N,$$

откуда вытекает условие согласования метода с порядком p .

Наоборот, пусть метод согласован с порядком p . Это означает, что $E_1(d_1 + \dots + d_l) = O(\Delta^p)$. Возьмем $t \in (\alpha, \beta]$ и рассмотрим такие сетки с шагом Δ и натуральные l , что $t - t_0 = l\Delta$. Тогда

$$E_1(d_1 + \dots + d_l) = \Delta^p E_1 \sum_{j=1}^l d^p(t_{j-1}) + O(\Delta^p) = \Delta^{p-1} E_1 \int_{t_0}^t d^p(s) ds + O(\Delta^p).$$

Но тогда получаем равенство

$$E_1 \int_{t_0}^t d^p(s) ds = 0,$$

продифференцировав которое, имеем $E_1 d^p(t) = 0$. \square

Теорема 5.3. Пусть метод (5.29) устойчив, функция Φ удовлетворяет условию квазилипищевости с порядком $p > 0$, оператор интерполирования I удовлетворяет условию квазилипищевости с порядком p , стартовые значения имеют порядок p , невязка обладает свойством гладкости с порядком p . Тогда метод сходится с порядком p тогда и только тогда, когда он согласован с порядком p .

Доказательство. По сравнению с доказательством теоремы 5.2, некоторые изменения произойдут во второй части. Равенство (5.30) заменяется на

$$S^{l-j} = E_1 + \lambda_2^{l-j} E_2 + \dots + \lambda_k^{l-j} E_k + \tilde{E}^{l-j}.$$

Тогда, с учетом леммы 5.4, получаем

$$\begin{aligned} \|R\|_S = \max_{0 \leq l \leq N} \left\| \sum_{j=0}^l S^{l-j} d_j \right\| &\leq \max_{0 \leq l \leq N} \left\| \Delta^p E_2 \sum_{j=1}^l \lambda_2^{l-j} d^p(t_{j-1}) + \dots + \right. \\ &\left. + \Delta^p E_k \sum_{j=1}^l \lambda_k^{l-j} d^p(t_{j-1}) + \sum_{j=0}^l \tilde{E}^{l-j} d_j \right\| + O(\Delta^p). \end{aligned}$$

Но в силу равенства

$$\sum_{j=1}^l \lambda^{l-j} d(t_{j-1}) = \frac{1 - \lambda^l}{1 - \lambda} d(t_0) + \sum_{j=1}^l \frac{1 - \lambda^{l-j}}{1 - \lambda} (d(t_j) - d(t_{j-1}))$$

выражение $\sum_{j=1}^l \lambda^{l-j} d(t_{j-1})$ равномерно ограничено, если $|\lambda| = 1$, $\lambda \neq 1$ и $d(t)$ имеет ограниченную вариацию.

Тогда с учетом того, что $\|\tilde{E}\| < 1$, выполняется $\|R\|_S = O(\Delta^p)$, что вместе с оценкой (5.12) леммы 5.1 показывает сходимость метода с порядком p . \square

5.5. Асимптотическое разложение глобальной погрешности

В этом разделе изучается асимптотическое разложение глобальной погрешности по величине шага дискретизации. Как известно, (см. [22, 40, 53]) в случае обыкновенных дифференциальных уравнений главный член разложения глобальной погрешности удовлетворяет некоторой линейной системе дифференциальных уравнений, сопряженной к исходной. В случае функционально-дифференциальных уравнений главный член разложения глобальной погрешности также удовлетворяет некоторой линейной системе, но уже функционально-дифференциальных уравнений. Асимптотическое разложение глобальной погрешности является теоретической основой для конструирования широкого класса экстраполяционных методов [22] и для организации процедур автоматического выбора шага. В первой части раздела рассматриваются одношаговые методы, во второй результаты обобщаются на произвольные сильноустойчивые методы.

Одношаговый метод. Рассмотрим одношаговый метод для задачи (5.1) – (5.2)

$$\begin{aligned} u_l &= \phi(t_l), \quad l = -m, \dots, 0, \\ u_{l+1} &= u_l + \Delta \Phi(t_l, u(\cdot), \Delta), \quad l = 0, \dots, N-1, \end{aligned} \tag{5.32}$$

с сеткой

$$t_0 = \alpha, \quad t_l = t_0 + l\Delta, \quad l = 0, \dots, N, \quad \Delta = \tau/m,$$

и с оператором интерполяции предыстории $u(\cdot) = I(\{u_i\}_{-m}^l)$ отображающим $\mathbb{R}^{n(l+m)}$ в пространства $X[\alpha - \tau, t_l + a\Delta]$, $a > 0$.

Одношаговый метод является частным случаем общей схемы раздела 3.2.2, в которой пространства дискретных моделей совпадают с фазовым пространством: $U_i = \mathbb{R}^n$, $l = -m, \dots, N$; функция точных значений задается фазовым вектором $Z(t_i, \Delta) = z_i = x(t_i)$, $i = -m, \dots, N$; стартовые значения модели совпадают с точными значениями $u_l = x(t_l)$, $l = -m, \dots, 0$; оператор S в пошаговой формуле задается единичной матрицей: $S = E$.

Например, это может быть явный или неявный метод типа Рунге–Кутты для системы (5.3)

$$\Phi(t_l, u(\cdot), \Delta) = \sum_{i=1}^k \sigma_i h_i,$$

$$h_i = h_i(t_l, u(\cdot)) = f(t_l + a_i\Delta, u_l + \Delta \sum_{j=1}^k b_{ij} h_j, u_{t_l + a_i\Delta}(\cdot)).$$

Сделаем дополнительные предположения, доопределив оператор интерполяции предыстории I и функцию продвижения модели на шаг Φ при $\Delta \rightarrow 0$.

Зафиксируем t_* и будем уменьшать шаг Δ , увеличивая l так, чтобы $t_l = t_0 + l\Delta = t_*$.

Пусть $x(t)$ – функция из множества $X[t_0 - \tau, t_*]$.

Условие 5.1. (Согласование предела оператора интерполяции предыстории I .)

$$I(\{x(t_i)\}_{-m}^l) \rightarrow x(t) \text{ при } \Delta \rightarrow 0 \text{ в пространстве } X[t_0 - \tau, t_*].$$

Условие 5.2. (Линейность оператора интерполяции предыстории I .)

Для всяких предысторий $\{u_i^1\}_{-m}^l$ и $\{u_i^2\}_{-m}^l$ и постоянных c^1 и c^2 выполняется

$$I(\{c^1 u_i^1 + c^2 u_i^2\}_{-m}^l) = c^1 I(\{u_i^1\}_{-m}^l) + c^2 I(\{u_i^2\}_{-m}^l).$$

Условие 5.3. (Ограниченность оператора интерполяции предыстории I .) Найдется постоянная $C = \|I\|$, такая, что для всяких предысторий $\{u_i\}_{-m}^l$ выполняется

$$\|I(\{u_i\}_{-m}^l)\| \leq C \|\{u_i\}_{-m}^l\|.$$

Заметим, что для оператора I интерполяции предыстории вырожденными сплайнами с экстраполяцией продолжением эти условия выполняются. Из ограниченности следует липшицевость оператора I , следовательно, квазилипшицевость с произвольным порядком p .

Обозначим (если предел существует)

$$\Phi(t_*, u(\cdot), 0) = \lim_{\Delta \rightarrow 0} \Phi(t_*, u(\cdot), \Delta).$$

Условие 5.4. (Согласование функции продвижения на шаг.)

Для точного решения системы (5.1) выполняется условие

$$\Phi(t_*, x_{t_*}(\cdot), 0) = f(t_*, x_{t_*}(\cdot)).$$

Заметим, что для методов типа Рунге–Кутты, если метод имеет порядок невязки $p > 0$, то согласно определению невязки это условие выполняется.

Условие 5.5. Функционал $\Phi(t_*, u(\cdot), \Delta)$ дифференцируем в смысле Фреше по второму аргументу, причем производная Фреше $\Phi'_{u(\cdot)}(t_*, u(\cdot), \Delta)$ непрерывна по Δ в нуле.

Для методов типа Рунге–Кутты дифференцируемость по Фреше функционала Φ следует из дифференцируемости по Фреше функционала f . Из дифференцируемости в смысле Фреше следует липшицевость функционала Φ по второму аргументу.

Теорема 5.4. Пусть погрешность аппроксимации метода (невязка) имеет разложение

$$d_l = \gamma(t) \Delta^{p+1} + o(\Delta^{p+1}), \quad l = 0, \dots, N-1, \quad (5.33)$$

где $\gamma(t)$ – непрерывная на $[t_0, t_0 + \theta]$ функция. Тогда при выполнении условий 5.1 – 5.5 глобальную погрешность метода можно представить в виде

$$u_l - x_l = e(t_l) \Delta^p + o(\Delta^p), \quad (5.34)$$

где функция $e(t)$ удовлетворяет на $[t_0, t_0 + \theta]$ уравнению

$$\dot{e}(t) = \langle G(t), e_t(\cdot) \rangle - \gamma(t) \quad (5.35)$$

с начальными условиями

$$e(t) = 0, \quad t \leq t_0.$$

Здесь $G(t) = \Phi'_{x(\cdot)}(t, x(\cdot), 0)$ – производная Фреше функционала $\Phi(t, x(\cdot), 0)$ в пространстве $X[t_0 - \tau, t]$, $\langle G(t), e_t(\cdot) \rangle$ – результат действия $G(t)$ на $e_t(\cdot) \in X[t_0 - \tau, t]$.

Доказательство. Рассмотрим величину

$$\hat{u}_l = u_l - e(t_l) \Delta^l, \quad l = 0, \dots, N,$$

как результат действия нового численного метода. Тогда для $l \geq 0$ получаем

$$\begin{aligned} \hat{u}_{l+1} &= u_{l+1} - e(t_{l+1}) \Delta^{l+1} = \\ &= \hat{u}_l + [e(t_l) - e(t_{l+1})] \Delta^p + \Delta \Phi(t_l, I(\{\hat{u}_l + e(t_l) \Delta^p\}_{-m}^l), \Delta) = \hat{u}_l + \Delta \hat{\Phi}, \end{aligned}$$

где

$$\begin{aligned} \hat{\Phi} &= \Phi(t_l, \hat{I}(\{\hat{u}_l\}_{-m}^l), \Delta) + [e(t_l) - e(t_{l+1})] \Delta^{p-1}, \\ \hat{I}(\{\hat{u}_l\}_{-m}^l) &= I(\{\hat{u}_l + e(t_l) \Delta^p\}_{-m}^l). \end{aligned}$$

Вычислим невязку этого нового метода

$$\hat{d}_l = x_{l+1} - x_l - \Delta \Phi(t_l, \hat{I}(\{x_l\}_{-m}^l), \Delta) - [e(t_l) - e(t_{l+1})] \Delta^p =$$

$$\begin{aligned}
&= \gamma(t_l)\Delta^{p+1} + \Delta[\Phi(t_l, I(\{x_i\}_{-m}^l), \Delta) - \\
&- \Phi(t_l, I(\{x_i + e(t_i)\Delta^p\}_{-m}^l), \Delta)] + [e(t_{l+1}) - e(t_l)]\Delta^p + o(\Delta^{p+1}) = \\
&= \gamma(t_l)\Delta^{p+1} - \Delta \langle \Psi'_{x(\cdot)}(t_l, I(\{x_i\}_{-m}^l), \Delta), I(\{e(t_i)\Delta^p\}_{-m}^l) \rangle + \\
&\quad + e'(t_l)\Delta^{p+1} + o(\Delta^{p+1}) = \\
&= \Delta^p[e'(t_l) - \langle \Psi'_{x(\cdot)}(t_l, x_{t_l}(\cdot), 0), e_{t_l}(\cdot) \rangle + \gamma(t_l)] + o(\Delta^p).
\end{aligned}$$

Если функция $e(t)$ удовлетворяет уравнению (5.35), то невязка будет иметь порядок выше чем p , (в смысле определения раздела 2.1.2). Так как стартовые значения имеют нулевую погрешность, оператор I липшицев, функция Φ липшицева по второму аргументу, то, согласно теореме 5.1, новый метод будет иметь порядок сходимости выше чем p . \square

Общая схема. Сильноустойчивые конечномерные методы. Рассмотрим случай общей схемы, в которой пространство моделей конечномерно: $U_l = \mathbb{R}^q$, $S_l = S \in \mathbb{R}^{q \times q}$, т.е. случай, рассматриваемый нами в разделе 3.2.4.

Подобно случаю одношаговых методов, сделаем дополнительные предположения, доопределив оператор интерполяции предыстории I и функцию продвижения модели на шаг Φ при $\Delta \rightarrow 0$.

Зафиксируем t_* и будем уменьшать шаг Δ , увеличивая l так, чтобы $t_l = t_0 + l\Delta = t_*$.

Пусть $e(t)$ — определенная на отрезке $[t_0 - \tau, t_*]$ q -мерная вектор-функция, принадлежащая некоторому линейному нормированному пространству $X^q[t_0 - \tau, t_*]$. При каждом l , таком, что $t_l = t_0 + l\Delta = t_*$ оператор интерполяции определяет элемент $v_l = I(e\{t_i\}_{-m}^l)$ интерполяционного пространства V_l . Пусть существует отображение Π пространств V_l в некоторое линейное нормированное пространство V_{t_*} . Определим значение оператора интерполяции на $e(t)$ соотношением

$$I(e(\cdot)) = \lim_{l \rightarrow \infty} \Pi I(e\{t_i\}_{-m}^l) \quad (5.36)$$

в предположении, что выполняется

Условие 5.6. Предел в (5.36) существует для функции точных значений $z(t) = \Lambda x(t)$ при $t \in [\alpha, \beta]$. (Доопределение оператора интерполяции.)

Условие 5.7. Оператор интерполяции предыстории I и его доопределение линейны.

Подобным же образом определим предельное значение функции продвижения модели на шаг $\Phi(t_l, v_l, \Delta)$ при таком стремлении Δ к нулю, что $t_l = t_0 + l\Delta = t_*$.

Пусть существует отображение Π пространств V_l в некоторое линейное нормированное пространство V_{t_*} , причем для всех последовательностей $v_l \in V_l$ существует $\lim_{l \rightarrow \infty} \Pi v_l = v^* \in V_{t_*}$. Определим $\Phi(t_*, v^*, 0)$ соотношением

$$\Phi(t_*, v^*, 0) = \lim_{l \rightarrow \infty} \Phi(t_l, v_l, \Delta) \quad (5.37)$$

в предположении, что выполняется

Условие 5.8. Предел в (5.37) существует. (Доопределение функции продвижения модели на шаг.)

Условие 5.9. Функционалы $\Phi(t_*, v, \Delta)$ дифференцируемы в смысле Фреше по второму аргументу, причем производные Фреше $\Phi'_v(t_*, v, \Delta)$ непрерывны по Δ в нуле.

Теорема 5.5. Пусть погрешность аппроксимации метода (невязка) имеет разложение

$$d_l = \gamma(t_l)\Delta^p + o(\Delta^p), \quad l = -l, \dots, 0, \quad (5.38)$$

$$d_l = d_p(t_l)\Delta^p + d_{p+1}(t_l)\Delta^{p+1} + o(\Delta^{p+1}), \quad l = 0, \dots, N, \quad (5.39)$$

где $\gamma(t)$ — непрерывная на $[t_0 - \tau, t_0]$ функция, $d_p(t)$ — непрерывно дифференцируема и $d_{p+1}(t)$ — непрерывна на $[t_0, \beta]$, причем метод сильно устойчив, оператор I квазилипицев с порядком выше чем p , функция Φ квазилипицева по второму аргументу с порядком выше чем p , и

$$E_1 d_p(t) = 0, \quad t \in [t_0, \beta],$$

(последнее условие следует из условия согласования с порядком p). Тогда при выполнении условий 5.6 – 5.9 глобальную погрешность метода можно представить в виде

$$u_l - x_l = (e(t_l) + \varepsilon_l)\Delta^p + o(\Delta^p), \quad (5.40)$$

где функция $e(t)$ удовлетворяет на $[t_0, \beta]$ линейному функционально-дифференциальному уравнению

$$\dot{e}(t) = E_1 \langle G(t), I(e_t(\cdot)) \rangle - E_1 d_{p+1}(t) - (E - S - E_1)^{-1} \dot{d}_p(t) \quad (5.41)$$

с начальными условиями

$$e(t_0) = -\gamma(t_0) - \varepsilon_0, \quad (5.42)$$

$$e(t) = -\gamma(t), \quad t < t_0. \quad (5.43)$$

Здесь $G(t) = \Phi'_v(t, I(z_t(\cdot)), 0)$ — производная Фреше функционала $\Phi(t, I(z_t(\cdot)), 0)$ по второму аргументу, величины ε_l , стремящиеся к нулю при $l \rightarrow \infty$, определяются соотношениями

$$\varepsilon_0 = (E_1 - E)\gamma(t_0) + (E - S + E_1)^{-1} d_p(t_0), \quad (5.44)$$

$$\varepsilon_l = (S - E_1)^l \varepsilon_0. \quad (5.45)$$

Доказательство. Рассмотрим величину

$$\hat{u}_l = u_l - e(t_l)\Delta^l, \quad l < 0,$$

$$\hat{u}_l = u_l - (e(t_l) + \varepsilon_l)\Delta^l, \quad l = 0, \dots, N,$$

как результат действия нового численного метода.

Тогда при $t < t_0$ невязка этого метода $\hat{d}_l = z_l - \hat{u}_l = (\gamma(t_l) + e(t_l))\Delta^p + o(\Delta^p)$ будет иметь порядок выше, чем p , если выполняется (5.43).

При $l = 0$ имеем

$$z_0 - \hat{u}_0 = z_0 - u_0 + (e(t_0) + \varepsilon_0)\Delta^p = (\gamma(t_0) + e(t_0) + \varepsilon_0)\Delta^p + o(\Delta^p). \quad (5.46)$$

При $l \geq 0$, согласно пошаговой формуле, получаем

$$\begin{aligned}\hat{u}_{l+1} &= u_{l+1} - (e(t_{l+1} + \varepsilon_{l+1})\Delta^l = \hat{u}_l + [Se(t_l) + S\varepsilon_l - e(t_{l+1}) - \varepsilon_{l+1}]\Delta^p + \\ &+ \Delta\Phi(t_l, I(\{\hat{u}_i + (e(t_i) + \varepsilon_i)\Delta^p\}_{-m}^l), \Delta) = S\hat{u}_l + \Delta\hat{\Phi},\end{aligned}$$

где

$$\begin{aligned}\hat{\Phi} &= \Phi(t_l, \hat{I}(\{\hat{u}_i\}_{-m}^l), \Delta) + [Se(t_l) + S\varepsilon_l - e(t_{l+1}) - \varepsilon_{l+1}]\Delta^{p-1}, \\ \hat{I}(\{\hat{u}_i\}_{-m}^l) &= I(\{\hat{u}_i + (e(t_i) + \varepsilon_i)\Delta^p\}_l).\end{aligned}$$

Вычислим невязку этого нового метода при $l \geq 0$

$$\begin{aligned}\hat{d}_l &= z_{l+1} - Sz_l - \Delta\Phi(t_l, \hat{I}(\{z_i\}_{-m}^l), \Delta) - [Se(t_l) + S\varepsilon_l - e(t_{l+1}) - \varepsilon_{l+1}]\Delta^p = \\ &= d_p(t_l)\Delta^p + d_{p+1}\Delta^{p+1} + \Delta\Phi(t_l, I(\{z_i\}_{-m}^l), \Delta) - \Delta\Phi(t_l, I(\{z_i + (e(t_i) + \\ &+ \varepsilon_i)\Delta^p\}_{-m}^l), \Delta) + [e(t_{l+1}) + \varepsilon_{l+1} - Se(t_l) - S\varepsilon_l]\Delta^p + o(\Delta^{p+1}).\end{aligned}\quad (5.47)$$

Из (5.46) – (5.47) следует, что невязка будет согласована с порядком выше чем p , если будут одновременно выполнены условия

$$\gamma(t_0) + e(t_0) + \varepsilon_0 = 0, \quad (5.48)$$

$$d_p(t_l) + e(t_l) - Se(t_l) + \varepsilon_{l+1} - S\varepsilon_l = 0, \quad (5.49)$$

$$E_1[\dot{e}(t_l) - \langle G(t), I(e_{t_l}(\cdot)) \rangle + d_{p+1}(t_l)] = 0, \quad (5.50)$$

$$E_1\varepsilon_l = 0. \quad (5.51)$$

Решим эту систему уравнений. Условие (5.49) будет выполняться, если одновременно выполнены условия

$$\varepsilon_{l+1} = S\varepsilon_l, \quad (5.52)$$

$$d_p(t) + e(t) - Se(t) = 0 \text{ при } t \geq t_0. \quad (5.53)$$

Условие (5.53) при $t = t_0$ примет вид

$$d_p(t_0) = -e(t_0) + Se(t_0). \quad (5.54)$$

Запишем (5.53) в эквивалентной форме:

$$d_p(t) = (S - E)e(t) \text{ при } t \geq t_0, \quad (5.55)$$

где E — единичная матрица.

Докажем соотношение

$$(E - S + E_1)^{-1}(E - S) = E - E_1. \quad (5.56)$$

В самом деле, из определения сильной устойчивости следует, что $S = E_1 + \tilde{E}$ (см. раздел 3.2.4). Отсюда $SE_1 = E_1^2 = E_1$. Справедливо очевидное тождество

$$(E - S + E_1)(E - E_1) = E - S,$$

откуда с учетом невырожденности матрицы $E - S - E_1$ (эта невырожденность вытекает из того, что $S - E_1$ не может иметь собственных чисел равных единице) следует (5.56).

Из (5.55) и (5.56) следует

$$(E - S + E_1)^{-1}d_p(t) = (E_1 - E)e(t). \quad (5.57)$$

Умножив (5.48) на $(E - E_1)$, получим с учетом (5.57)

$$(E - E_1)\varepsilon_0 = (E_1 - E)\gamma(t_0) + (E - S + E_1)^{-1}d_p(t_0). \quad (5.58)$$

Выберем ε_0 из соотношения (5.44). Тогда, из условия согласования с порядком p следует $E_1\gamma(t_0) = 0$ и $E_1d_p(t_0) = 0$, поэтому

$$E_1\varepsilon_0 = 0, \quad (5.59)$$

и соотношение (5.58), а, следовательно, и соотношение (5.48) выполнится.

Выберем ε_l из соотношения (5.52) (автоматически выполнится условие (5.51)). Тогда, из условия (5.59) вытекает

$$\varepsilon_l = S^l\varepsilon_0 = (E_1 + S - E_1)^l\varepsilon_0 = (E_1 - E)\gamma(t_0) + (E - S + E_1)^{-1}d_p(t_0),$$

и из сильной устойчивости следует, что величины ε_l стремятся к нулю при $l \rightarrow \infty$.

Выберем теперь величину $e(t)$ как решение уравнения (5.50) с начальными условиями (5.42) – (5.43).

Введем величину $v(t) = E_1e(t)$, тогда, в силу (5.57),

$$e(t) = v(t) - (E - S + E_1)^{-1}d_p(t).$$

Дифференцируя это соотношение, получаем, в силу (5.50), уравнение (5.41). \square

Замечание 5.2. Случай метода со слабо устойчивой матрицей S проводится аналогично [40].

Замечание 5.3. Подобно [22, 40] можно выписать уравнения для следующих членов глобальной погрешности при более сильных предположениях гладкости функционала Φ и некоторых дополнительных предположениях относительно стартовых значений, оператора интерполяции и коэффициентов разложения в невязке.

Глава 6.

Алгоритмы с переменным шагом и некоторые вопросы компьютерной реализации численных моделей

В предыдущих разделах рассматривались численные методы решения ФДУ с постоянным шагом дискретизации Δ . Однако современные программные средства численного решения ОДУ используют процедуры с автоматическим выбором шага в зависимости от поведения решения системы. Некоторые модификации приведенных выше главах алгоритмов позволяют перенести большую часть результатов и на случай переменного шага. В качестве примера такой модификации приводится описание алгоритмов и доказываются соответствующие теоремы сходимости для явных методов типа Рунге–Кутты с интерполяцией вырожденными сплайнами и экстраполяцией продолжением.

Эти алгоритмы, наряду с процедурой автоматического выбора шага (с помощью оценки главного члена локальной погрешности вложенных методов), послужили основой для создания соответствующего программного обеспечения в виде пакета программ Time-Delay System Toolbox [45], предназначенного для численного решения широкого класса систем с постоянным, переменным и распределенным запаздыванием и для моделирования некоторых задач управления такими системами.

Нами приводится также описание некоторых тестовых задач, на которых проводилось компьютерное экспериментирование при реализации приведенных численных методов и описание некоторых модельных задач, при решении которых применялись эти методы.

6.1. ЯРК-методы с переменным шагом

Рассмотрим как и в разделе 3.1 систему ФДУ

$$\dot{x}(t) = f(t, x(t), x_t(\cdot)), \quad (6.1)$$

с начальными условиями

$$x(t_0) = x_0, \quad (6.2)$$

$$x_{t_0}(\cdot) = \{y^0(s), -\tau \leq s < 0\}, \quad (6.3)$$

с теми же предположениями относительно функционала f .

Зададим на $[t_0, t_0 + \theta]$ неравномерную *временную сетку* $t_{l+1} = t_l + \Delta_l$, $l = 0, 1, \dots, N-1$, где $\Delta_l > 0$ — шаг. Обозначим $\Delta_{\max} = \max_l \Delta_l$, $\Delta_{\min} = \min_l \Delta_l$. Будем предполагать, что $\Delta_{\max} \leq K \Delta_{\min}$, K — фиксированное число.

Введем *дискретную численную модель* системы (6.1), обозначив приближение решения $x(t_l) = x_l$ в точке t_l через $u_l \in \mathbb{R}^\ell$.

Специфика системы (6.1) состоит в том, что адекватная ей дискретная модель должна при построении учитывать в момент t_l *предысторию дискретной модели*, т.е. множество $\{u_i\}_l = \{u_i, t_i \in$

$\in [t_l - \tau, t_l]$. Для методов с переменным шагом нужно учитывать более длинную предысторию.

Определение 6.1. *Расширенной предысторией дискретной модели назовем множество $\{u_i\}_l^+ = \{u_i, t_i \in [t_l - 2\tau, t_l]\}$.*

Определение 6.2. *Пусть $a > 0$. Будем говорить, что на множестве расширенных дискретных предысторий модели $\{u_i\}_l^+$ задан оператор интерполяции и экстраполяции IE , если*

$$IE(\{u_i\}_l^+) = u(\cdot) \in Q[t_l - \tau, t_l + a\Delta_{max}]. \quad (6.4)$$

В дальнейшем будем предполагать, что

1) оператор IE согласован, т.е.

$$u(t_i) = u_i, \quad t_i \in [t_l - \tau, t_l]; \quad (6.5)$$

2) оператор IE удовлетворяет условию Липшица, т.е. найдется константа L_{IE} такая, что для всяких двух наборов дискретных предысторий $\{u_i^1\}_l^+$ и $\{u_i^2\}_l^+$ и для всех $t \in [t_l - \tau, t_l + a\Delta_{max}]$ выполняется условие

$$\|u^1(t) - u^2(t)\| \leq L_{IE} \max_{t_l - 2\tau \leq t_i \leq t_l} \|u_i^1 - u_i^2\|, \quad (6.6)$$

где $u^1(\cdot) = IE(\{u_i^1\}_l^+)$, $u^2(\cdot) = IE(\{u_i^2\}_l^+)$.

Определение 6.3. *Будем говорить, что оператор интерполяции и экстраполяции IE имеет порядок p на некотором множестве функций $X \subseteq Q[t_l - \tau, t_l + a\Delta_{max}]$, если найдется константа C_{IE} , такая, что для всяких $x(\cdot) \in X$ и $t \in [t_l - \tau, t_l + a\Delta_{max}]$ выполняется*

$$\|x(t) - \tilde{x}(t)\| \leq C_{IE}\Delta^p, \quad (6.7)$$

где $\tilde{x}(\cdot) = IE(\{x_i\}_l^+)$, $x_i = x(t_i)$, $t_i \in [t_l - 2\tau, t_l]$.

Так, оператор кусочно-постоянной интерполяции и экстраполяции $IE: \{u_i\}_l \rightarrow u(\cdot)$

$$u(t) = \begin{cases} u_i, & t \in [t_i, t_{i+1}), \\ u_l, & t \in [t_l, t_l + a\Delta_{max}) \end{cases}$$

согласован, удовлетворяет условию Липшица с $L_{IE} = 1$ и на множестве липшицевых функций $x(t)$ имеет первый порядок. Примеры операторов интерполяции и экстраполяции более высокого порядка приведем в разделе 3.3.2.

Определение 6.4. *Явным k -этапным методом типа Рунге-Кутты (ЯРК) с переменным шагом назовем дискретную модель вида*

$$u_0 = x_0, \quad u_{l+1} = u_l + \Delta_l \Psi(t_l, u(\cdot)), \quad l = 1, \dots, N-1, \quad (6.8)$$

$$u(\cdot) = IE(\{u_i\}_l^+), \quad \Psi(t_l, u(\cdot)) = \sum_{i=1}^k \sigma_i h_i, \quad (6.9)$$

$$h_1 = h_1(t_l, u(\cdot)) = f(t_l, u_l, u_l(\cdot)), \quad (6.10)$$

$$h_i = h_i(t_l, u(\cdot)) = f(t_l + a_i \Delta_l, u_l + \Delta_l \sum_{j=1}^{i-1} b_{ij} h_j, u_{t_l + a_i \Delta_l}(\cdot)), \quad (6.11)$$

$$|a_i| \leq a, \quad i = 1, \dots, k.$$

Коэффициенты метода удобно задаются таблицей Бутчера

0	0	0	...	0
a_2	b_{21}	0	...	0
.
a_k	b_{k1}	b_{k2}	...	0
	σ_1	σ_2	...	σ_k

Будем обозначать $a = \max\{|a_i|, 1 \leq i \leq k\}$, $\sigma = \max\{|\sigma_i|\}$, $b = \max\{|b_{ij}|\}$.

Справедливо утверждение

Лемма 6.1. *Функционалы h_1, h_2, \dots, h_k , а также Ψ липшицевы по $u(\cdot)$.*

Доказательство этой леммы повторяет доказательство соответствующего утверждения (лемма 3.4) из раздела 3.1.3.

Следствие 6.1. *Если оператор интерполяции-экстраполяции IE задан и удовлетворяет условию Липшица с константой L_{IE} , то функция $\Phi(\{u_i\}_l^+) = \Psi(t_l, IE(\{u_i\}_l^+))$ липшицева по расширенной дискретной предыстории модели, т.е. найдется такая константа $L_\Phi = L_\Psi L_{IE}$, при которой*

$$\|\Phi(\{u_i^1\}_l^+) - \Phi(\{u_i^2\}_l^+)\| \leq L_\Phi \max_{t_l - 2\tau \leq t_i \leq t_l} \|u_i^1 - u_i^2\|.$$

Пусть $x_l = x(t_l)$, $l = 0, 1, \dots, N$ — значения точного решения $x(t)$ задачи (6.1) — (6.3) в узлах t_l .

Невязкой ЯРК-метода на точном решении $x(\cdot) = \{x(t), t \in [t_0 - \tau, t_0 + \theta]\}$ в узлах t_l назовем функционалы

$$g_l(x(\cdot)) = \frac{x_{l+1} - x_l}{\Delta_l} - \Psi(t_l, x(\cdot)), \quad (6.12)$$

где

$$\Psi(t_l, x(\cdot)) = \sum_{i=1}^{i-1} \sigma_i h_i, \quad h_i = f(t_l + a_i \Delta_l, x_l + \Delta_l \sum_{j=1}^{i-1} b_{ij} h_j, x_{t_l + a_i \Delta_l}(\cdot)).$$

В ОДУ порядок невязки определяется обычно на некотором классе задач, например на классе достаточно гладких задач. В ФДУ задача (6.1) — (6.3) задается тройкой (f, x_0, y^0) , т.е. функционалом правой части системы (6.1), конечномерной x_0 и бесконечномерной y^0 составляющими в начальных условиях. Пусть $F = \{(f, x_0, y^0)\}$ — некоторое подмножество (класс задач) таких троек, которые удовлетворяют указанным в разделе 1.4.1 условиям, и пусть $X(F)$ — соответствующее множество решений $x(t)$ задач (6.1) — (6.3).

Будем говорить, что *ЯРК-метод имеет порядок аппроксимации (невязки) p на классе задач F* , если для всякого $x(\cdot) \in X(F)$ найдется константа $C_{x(\cdot)}$ такая, что

$$\|g_l(x(\cdot))\| \leq C_{x(\cdot)} \Delta_{max}^p, \quad l = 0, \dots, N-1. \quad (6.13)$$

Будем говорить, что ЯРК-метод (6.8) – (6.11) для задачи (6.1) – (6.3) *сходится*, если

$$\max_{1 \leq l \leq N} \|u_l - x(t_l)\| \rightarrow 0 \text{ при } \Delta_{max} \rightarrow 0,$$

и метод *сходится с порядком p* , если найдется такая константа C_u , что

$$\|x_l - u_l\| \leq C_u \Delta_{max}^p, \quad l = 0, 1, \dots, N \quad (6.14)$$

В ОДУ порядок сходимости ЯРК-метода определяется лишь порядком аппроксимации, в ФДУ он зависит также от порядка оператора интерполяции и экстраполяции, а именно справедлива

Теорема 6.1. Пусть ЯРК-метод (6.8) – (6.11) имеет порядок аппроксимации $p_1 > 0$ и оператор интерполяции-экстраполяции IE имеет порядок $p_2 > 0$. Тогда ЯРК-метод сходится и имеет порядок сходимости $p = \min\{p_1, p_2\}$.

Доказательство. Обозначим $\delta_l = u_l - x_l$, тогда из (6.8) следует $\delta_0 = 0$. Для $l = 0, \dots, N-1$ имеем

$$\begin{aligned} \delta_{l+1} &= \delta_l + \Delta_l [\Psi(t_l, IE(\{u_i\}_l^+)) - \Psi(t_l, IE(\{x_i\}_l^+))] + \\ &+ [x_l - x_{l+1} + \Delta_l \Psi(t_l, x(\cdot))] + \Delta_l [\Psi(t_l, I(\{x_i\}_l^+)) - \Psi(t_l, x(\cdot))]. \end{aligned}$$

Используя для функции $\Phi(\{u_i\}_l^+) = \Psi(t_l, IE(\{u_i\}_l^+))$ следствие 6.1, учитывая (6.13) и (6.11), получим

$$\|\delta_{l+1}\| \leq \|\delta_l\| + \Delta_l L_\Phi \max_{t_l - 2\tau \leq t_i \leq t_l} \|\delta_i\| + C_{x(\cdot)} \Delta_l^{p+1} + L_\Psi C_{IE} \Delta_l^{p+1}. \quad (6.15)$$

Индукцией по l докажем оценку

$$\|\delta_l\| \leq (1 + \Delta_{max}(L_\Phi + 1))^l (C_{x(\cdot)} + L_\Psi C_{IE}) \Delta_{max}^p, \quad l = 0, \dots, N. \quad (6.16)$$

База индукции выполняется в силу $\delta_0 = 0$.

Шаг индукции. Предположим, что оценка (6.16) верна для всех индексов $\leq l$. Докажем ее для $l+1$.

В силу (6.15) и индуктивного предположения

$$\begin{aligned} \|\delta_{l+1}\| &\leq (1 + \Delta_{max}(L_\Phi + 1))^l (C_{x(\cdot)} + L_\Psi C_{IE}) \Delta_{max}^p + \\ &+ \Delta_{max} L_\Phi (1 + \Delta_{max}(L_\Phi + 1))^{l_0} (C_{x(\cdot)} + L_\Psi C_{IE}) \Delta_{max}^p + (C_{x(\cdot)} + L_\Psi C_{IE}) \Delta_{max}^{p+1}, \end{aligned}$$

где $l_0 \leq l$ — индекс, на котором достигается $\max_{t_l - 2\tau \leq t_i \leq t_l} \|\delta_i\|$. Тогда

$$\|\delta_{l+1}\| \leq (1 + \Delta_{max}(L_\Phi + 1))^l (1 + \Delta_{max} L_\Phi + \Delta_{max}) (C_{x(\cdot)} + L_\Psi C_{IE}) \Delta_{max}^p,$$

т.е. получаем оценку (6.16) для $l+1$.

Так как число переменных шагов

$$N \leq \frac{\theta}{\Delta_{\min}} \leq \frac{K\theta}{\Delta_{\max}},$$

то из (6.16) вытекает оценка

$$\begin{aligned} \|x_l - u_l\| &\leq (1 + \Delta_{\max}(L_\Phi + 1))^N (C_{x(\cdot)} + L_\Psi C_{IE}) \Delta_{\max}^p \leq \\ &\leq (C_{x(\cdot)} + L_\Psi C_{IE}) e^{K(L_\Phi + 1)\theta} \Delta_{\max}^p, \quad l = 0, 1, \dots, N, \end{aligned}$$

которая содержит утверждение теоремы. \square

Как видно, основным моментом в конструировании методов высокого порядка точности служит (наряду с качествами оператора интерполяции-экстраполяции) эффективное определение порядка невязки, и следовательно, подбор коэффициентов в матрице Бутчера. Для определения порядка невязки все результаты изложенные в разделе 3.1.4 без изменений переносятся на случай переменного шага.

6.2. Способы интерполяции и экстраполяции расширенной предыстории дискретной модели

Пусть на отрезке $[t_0, t_0 + \theta]$ задана временная сетка, такая же, как и в разделе 3.3.1.

Оператором *интерполирования I расширенной дискретной предыстории модели* назовем отображение $I : \{u_i\}_l^+ \rightarrow u(\cdot) \in Q[t_l - \tau, t_l]$.

Будем говорить, что оператор I имеет порядок погрешности p на точном решении, если существуют константы C_1, C_2 такие, что для всех $l = 0, 1, \dots, N$ и $t \in [t_l - \tau, t_l]$ выполняется неравенство

$$\|x(t) - u(t)\| \leq C_1 \max_{i \geq 0, t_l - 2\tau \leq t_i \leq t_l} \|u_i - x_i\| + C_2 \Delta_{\max}^p. \quad (6.17)$$

Примером оператора интерполяции второго порядка на достаточно гладком точном решении служит интерполяция ломаными (кусочно-линейная интерполяция):

$$\begin{aligned} I : \{u_i\}_l^+ &\rightarrow u(t) = \\ &= \begin{cases} ((t - t_i)u_{i+1} + (t_{i+1} - t)u_i) \frac{1}{\Delta_i}, & t \in [t_i, t_{i+1}], \quad t_{i+1} \in [t_l - \tau, t_l], \\ y^0(t_0 - t), & t \in [t_l - \tau, t_0]. \end{cases} \end{aligned} \quad (6.18)$$

Отметим, что если $t \in [t_i, t_{i+1}]$, $t_{i+1} \in [t_l - \tau, t_l]$, но $t_i \notin [t_l - \tau, t_l]$, то в построении $u(t)$ существенно участвует расширенная предыстория. Такие ситуации исключаются в интерполяции на равномерной временной сетке, где предполагается, что шаг сетки Δ кратен величине τ запаздывания (см. раздел 3.1.2).

Рассмотрим более общий способ интерполяции функциями, кусочно составленными из многочленов p -й степени, где p — произвольное натуральное число. Предположим, что $\tau/\Delta_{\max} \geq p$, т.е. максимальный шаг достаточно мал.

Разобьем отрезок $[\max\{t_0, t_l - \tau\}, t_l]$ справа на подотрезки $[t_{l_{i+1}}, t_{l_i}]$, $i = 0, 1, \dots, k-1$, объединяя $p+1$ соседних узлов сетки: $t_{l_0} = t_l$, $t_{l_1} = t_{l-p}, \dots, t_{l-ip}, \dots$. Последний подотрезок $[t_{l_k}, t_{l_{k-1}}]$ может содержать меньше чем $p+1$ узлов t_i сетки из отрезка $[\min\{t_0, t_l - \tau\}, t_l]$, поэтому добавим в него узлы t_i из отрезка $[t_l - 2\tau, t_l - \tau)$. Если

узлов не хватает в силу $t_i < t_0$, то построим фиктивные узлы $t_i = t_0 + i\Delta_0$, $i < 0$, значения в которых возьмем равными $u_i = y^0(t_i - t_0)$.

На каждом отрезке $[t_{l_{i+1}}, t_{l_i}]$ построим интерполяционный многочлен $L_p(t) = L_p^i(t)$ по данным $u_{l_i-p}, u_{l_i-p+1}, \dots, u_{l_i}$:

$$L_p^i(t) = \sum_{n=0}^p u_{l_i-n} \prod_{j=l_i-p; j \neq l_i-n}^{l_i} \frac{t - t_j}{t_{l_i-n} - t_j}.$$

Определим оператор интерполирования предыстории модели

$$I : \{u_i\}_l \rightarrow u(t) = \begin{cases} L_p^i(t), & t_{l_{i+1}} \leq t < t_{l_i}, \quad i = 0, \dots, l-1, \\ y^0(t_0 - t), & t \in [t_l - \tau, t_0). \end{cases} \quad (6.19)$$

Теорема 6.2. Пусть точное решение $x(t)$ является $p+1$ раз непрерывно дифференцируемо на отрезке $[t_0 - \tau, t_0 + \theta]$, тогда оператор интерполяции вырожденными сплайнами p -го порядка имеет погрешность интерполяции $p+1$.

Доказательство. Возьмем произвольные $l = 1, \dots, N$ и $t \in [\max\{t_0, t_l - \tau\}, t_l]$. Пусть t лежит в подотрезке $[t_{l_{i+1}}, t_{l_i}]$. Тогда

$$\|x(t) - u(t)\| = \|x(t) - L_p^i(t)\| \leq \|x(t) - \hat{L}_p^i(t)\| + \|L_p^i(t) - \hat{L}_p^i(t)\|,$$

где $\hat{L}_p^i(t)$ — интерполяционный многочлен p -й степени, построенный по значениям x_j на отрезке $[t_{l_{i+1}}, t_{l_i}]$:

$$\hat{L}_p^i(t) = \sum_{n=0}^p x_{l_i-n} \prod_{j=l_{i+1}; j \neq l_i-n}^{l_i} \frac{t - t_j}{t_{l_i-n} - t_j}.$$

Погрешность интерполяции точного решения вычисляется по формуле [19, с. 133]

$$\|R_p(t)\| = \|x(t) - \hat{L}_p^i(t)\| = \left\| \frac{x^{(p+1)}(\xi)}{(p+1)!} \prod_{j=0}^p (t - t_{l_i-j}) \right\|,$$

где $\xi \in [t_{l_{i+1}}, t_{l_i}]$. Таким образом, если $t \in [t_{l_{i+1}}, t_{l_i}]$, то

$$\left\| \prod_{j=0}^p (t - t_{l_i-j}) \right\| \leq \Delta_{\max} \cdot \Delta_{\max} \cdot (2\Delta_{\max}) \cdots (p\Delta_{\max}) = \Delta_{\max}^{p+1} p!$$

и

$$\|R_p(t)\| \leq \frac{M_{p+1}}{(p+1)!} p! \Delta_{\max}^{p+1} = \frac{M_{p+1}}{p+1} \Delta_{\max}^{p+1}, \quad (6.20)$$

где $M_{p+1} = \max_{t_0 - \tau \leq t \leq t_0 + \theta} \|x^{(p+1)}(t)\|$. Кроме того, имеем оценку

$$\begin{aligned} \|L_p^i(t) - \hat{L}_p^i(t)\| &= \sum_{n=0}^p \|u_{l_i-n} - x_{l_i-n}\| \prod_{j=l_i-p; j \neq l_i-n}^{l_i} \frac{|t - t_j|}{|t_{l_i-n} - t_j|} \leq \\ &\leq \max_{l_i-p \leq n \leq l_i} \|u_n - x_n\| p! \frac{\Delta_{\max}}{\Delta_{\min}}. \end{aligned} \quad (6.21)$$

Из (6.20) и (6.21) получаем

$$\|x(t) - L_p^i(t)\| \leq C_1 \max_{l_i - p \leq n \leq l_i} \|u_n - x_n\| + C_2 \Delta^{p+1}. \quad (6.22)$$

Здесь $C_1 = Kp!$, $C_2 = \frac{M_{p+1}}{p+1}$. □

Кроме интерполяции вида (6.19) можно применять интерполяцию многочленами Эрмита с кратными узлами, используя условие $\dot{u}_l = f(t_l, u_l, u_{t_l}(\cdot))$, а также другие способы интерполяции дискретной предыстории модели.

В НРК-методах необходимо в момент t_l знать предысторию модели $u_{t_l+a}(\cdot)$ при $a > 0$, т.е. произвести экстраполяцию модели на отрезок $[t_l, t_l + a]$.

Для любого $a > 0$ оператором *экстраполирования* E предыстории модели назовем отображение $E : \{u_i\}_l^+ \rightarrow u(\cdot) \in Q[t_l, t_l + a\Delta_l]$.

Будем говорить, что экстраполяция предыстории модели имеет порядок погрешности p на точном решении, если существуют константы C_3, C_4 такие, что для всех $a > 0$ и всех $l = 0, 1, \dots, N-1$ и $t \in [t_l, t_l + a\Delta_l]$ выполняется неравенство

$$\|x(t) - u(t)\| \leq C_3 \max_{t_l - 2\tau \leq t_i \leq t_l} \|u_i - x_i\| + C_4 (\Delta_{max})^p. \quad (6.23)$$

Один из способов задания оператора экстраполяции — экстраполяция *продолжением* интерполяционного многочлена:

$$E : \{u_i\}_l \rightarrow u(t) = L_p^0(t), \quad t \in [t_l, t_l + a\Delta_l], \quad (6.24)$$

где $L_p^0(t)$ — интерполяционный многочлен p -ой степени, построенный по значениям u_j на отрезке $[t_{l-p}, t_l]$:

$$L_p^0(t) = \sum_{n=0}^p u_{l-n} \prod_{j=l-p; j \neq l-n}^l \frac{t - t_j}{t_{l-n} - t_j}.$$

Теорема 6.3. Пусть точное решение $x(t)$ является $p+1$ раз непрерывно дифференцируемо на отрезке $[t_0 - \tau, t_0 + \theta]$, тогда оператор экстраполяции продолжением интерполяционного многочлена p -й степени имеет погрешность экстраполяции порядка $p+1$.

Доказательство. Возьмем произвольное $t \in [t_l, t_l + a\Delta]$. Тогда

$$\|x(t) - u(t)\| = \|x(t) - L_p^0(t)\| \leq \|x(t) - \hat{L}_p^0(t)\| + \|L_p^0(t) - \hat{L}_p^0(t)\|,$$

где $\hat{L}_p^0(t)$ — интерполяционный многочлен p -й степени, построенный по значениям x_j на отрезке $[t_{l-p}, t_l]$.

Тогда

$$\begin{aligned} \|x(t) - \hat{L}_p^0(t)\| &\leq \frac{M_{p+1}}{(p+1)!} \prod_{j=0}^p (t - t_{l-j}) \leq \\ &\leq \frac{M_{p+1}}{(p+1)!} \prod_{j=0}^p (a\Delta_{max} + j\Delta_{max}) \leq C_4 (\Delta)_{max}^{p+1}, \end{aligned}$$

где

$$C_4 = \frac{M_{p+1}}{(p+1)!} \prod_{j=0}^p (a+j).$$

Кроме того,

$$\begin{aligned} \|L_p^0(t) - \hat{L}_p^0(t)\| &\leq \max_{l-p \leq n \leq l} \|u_n - x_n\| \sum_{n=0}^p \prod_{j=0; j \neq n}^p \frac{(a\Delta_{max} + j\Delta_{max})}{\Delta_{min}} \leq \\ &\leq C_3 \max_{l-p \leq n \leq l} \|u_n - x_n\|. \end{aligned}$$

Здесь

$$C_3 = (p+1) \prod_{j=0}^p (a+j).$$

Теорема доказана. \square

Объединяя утверждения теорем 6.1 и 6.2 и учитывая определения оператора интерполяции и экстраполяции IE и его порядка, получаем

Следствие 6.2. *Построенный оператор интерполяции вырожденными сплайнами порядка p и экстраполяции продолжением имеет порядок $p+1$ на достаточно гладком решении.*

6.2.1. Выбор длины шага

Приведём *правило Рунге практической оценки погрешности* применительно к рассматриваемым методам.

Предположим, что, пользуясь некоторым методом порядка p , имеем при переходе от нулевого шага к первому асимптотическое представление погрешности

$$\varepsilon_1 = x_1 - u_1 = C\Delta^{p+1} + O(\Delta^{p+2}).$$

Сделаем два шага длиной Δ и один длиной 2Δ . Погрешность при двух шагах состоит из перенесённой погрешности первого шага и локальной погрешности второго, и тогда она выражается как

$$\varepsilon_2 = x_2 - u_2 = x_2 - \hat{u}_2 + \hat{u}_2 - u_2,$$

где \hat{u}_2 — численное решение, полученное в момент t_2 из позиции $\{t_1, x(t_1), x_{t_1}(\cdot)\}$ как из начальной. Тогда

$$\hat{u}_2 - u_2 = x_1 - u_1 + \Delta(\Psi(IE(\{x_i\}_1)) - \Psi(IE(\{u_i\}_1))).$$

В силу леммы 6.1 и её следствия выполняется условие

$$\|\Psi(IE(\{x_i\}_1)) - \Psi(IE(\{u_i\}_1))\| \leq L_\Phi |x_1 - u_1|,$$

откуда

$$\begin{aligned} \varepsilon_2 &= C\Delta^{p+1} + C\Delta^{p+1}(1 + O(\Delta)) + O(\Delta^{p+2}) = \\ &= 2C\Delta^{p+1} + O(\Delta^{p+2}). \end{aligned}$$

Для значения w , полученного применением этого же метода с шагом 2Δ , имеем

$$x_2 - w = C(2\Delta)^{p+1} + O(\Delta^{p+2}).$$

Вычитая одно соотношение из другого, имеем

$$u_2 - w \approx 2C\Delta^{p+1}(2^p - 1),$$

откуда

$$\varepsilon_2 \approx \frac{u_2 - w}{2^p - 1}.$$

Эта формула не только дает оценку погрешности, но и возможность уточнить результат:

$$\hat{u}_2 = u_2 + \frac{u_2 - w}{2^p - 1}.$$

Такая процедура называется *экстраполяцией по Ричардсону*. На её основе для ОДУ создан целый класс *экстраполяционных методов*, среди которых самым мощным является алгоритм Грэгга-Булирша-Штёра (ГБШ) [22].

На основе приближенной оценки погрешности

$$\varepsilon \approx \frac{\|u_2 - w\|}{2^p - 1}$$

можно организовать процедуру автоматического выбора длины шага, обеспечивающую заданную допустимую точность tol . Погрешность будем обозначать err , как это принято в стандартных программах.

Если выбрана начальная величина шага Δ_{old} , то программа выполняет вычисление двух шагов этой длины и одного шага двойной длины. Вычисляется погрешность

$$err = \frac{1}{2^p - 1} \max_{i=1, \dots, n} \frac{|u_{2,i} - w_i|}{d_i},$$

где индекс i обозначает координату вектора, d_i — масштабирующий множитель. При $d_i = 1$ получаем абсолютную погрешность, при $d_i = |u_{2,i}|$ — относительную. Часто применяют и другие нормы и масштабирования.

Из соотношений

$$\begin{aligned} err &= C(2\Delta_{old})^{p+1}, \\ tol &= C(2\Delta_{new})^{p+1} \end{aligned}$$

получаем выражение для нового шага

$$\Delta_{new} = \left(\frac{tol}{err} \right)^{\frac{1}{p+1}} \Delta_{old}.$$

Возможны две ситуации:

- 1) $\Delta_{new} < \Delta_{old}$. Тогда старый шаг не принимается, он уменьшается до величины Δ_{new} .
- 2) $\Delta_{new} > \Delta_{old}$. Тогда проделанные два старых шага принимаются, а следующий шаг делается новым, возможно большим.

При практической реализации этой процедуры обычно не допускают слишком резкого увеличения или уменьшения шага, т.е. формула для нового шага более сложная:

$$\Delta_{new} = \min \left\{ fact_{max}, \max \left\{ fact_{min}, fac \left(\frac{tol}{err} \right)^{\frac{1}{p+1}} \right\} \right\} \Delta_{old},$$

где величину fac берут равной обычно 0.8, а величины $fact_{max}$ и $fact_{min}$ задают максимальный и минимальный коэффициенты увеличения и уменьшения шага.

Кроме того, при выборе шага обязательно должно соблюдаться условие

$$\Delta_{min} \leq \Delta_{new} \leq \Delta_{max}.$$

Идея *вложенных формул* Рунге–Кутты состоит в том, чтобы для оценки погрешности и автоматического выбора шага использовать не два разных шага по одному и тому же методу, а один шаг, но по двум методам разного порядка. Особая эффективность достигается, если коэффициенты a_i , b_{ij} в матрице Бутчера, а следовательно и h_i , для двух методов совпадают (такие методы называются *вложенными*). Метод p -го порядка

$$u_{l+1} = u_l + \Delta_l \sum_{i=1}^k \sigma_i h_i(t_l, u(\cdot))$$

является основным, а метод $p + 1$ -го порядка

$$\hat{u}_{l+1} = u_l + \Delta_l \sum_{i=1}^k \hat{\sigma}_i h_i(t_l, u(\cdot))$$

используется лишь для оценки погрешности.

Простым примером вложенных методов является пара — метод Эйлера с пересчетом и Рунге–Кутты третьего порядка:

0			
1	1		
$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$	
u_{l+1}	$\frac{1}{2}$	$\frac{1}{2}$	0
\hat{u}_{l+1}	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{4}{6}$

Этот метод называется методом *Рунге–Кутты–Фельберга второго-третьего порядка* и обозначается РКФ 2(3).

Другой, более точный, шестиэтапный метод *Рунге–Кутты–Фельберга четвертого-пятого порядка* РКФ 4(5)

0						
$\frac{1}{4}$	$\frac{1}{4}$					
$\frac{3}{8}$	$\frac{3}{32}$	$\frac{9}{32}$				
$\frac{12}{13}$	$\frac{1932}{2197}$	$-\frac{7200}{2197}$	$\frac{7296}{2197}$			
1	$\frac{439}{216}$	-8	$\frac{3680}{513}$	$-\frac{845}{4104}$		
$\frac{1}{2}$	$-\frac{8}{27}$	2	$-\frac{3544}{2565}$	$\frac{1859}{4104}$	$-\frac{11}{40}$	
u_{l+1}	$\frac{25}{216}$	0	$\frac{1408}{2565}$	$\frac{2197}{4104}$	$-\frac{1}{5}$	0
\hat{u}_{l+1}	$\frac{16}{135}$	0	$\frac{6656}{12825}$	$\frac{28561}{56430}$	$-\frac{9}{50}$	$\frac{2}{55}$

используется как основной при решении ОДУ в большинстве пакетов прикладных программ. Для решения ФДУ его необходимо дополнить интерполяцией и экстраполяцией пятого порядка.

Практически ту же точность имеет вложенный семиэтапный *метод Дормана–Принса 5(4)*:

0							
$\frac{1}{5}$	$\frac{1}{5}$						
$\frac{3}{10}$	$\frac{3}{40}$	$\frac{9}{40}$					
$\frac{4}{5}$	$\frac{44}{45}$	$-\frac{56}{15}$	$\frac{32}{9}$				
$\frac{8}{9}$	$\frac{19372}{6561}$	$-\frac{25360}{2187}$	$\frac{64448}{6561}$	$-\frac{212}{729}$			
1	$\frac{9017}{3168}$	$-\frac{355}{33}$	$\frac{46732}{5247}$	$\frac{49}{176}$	$-\frac{5103}{18656}$		
1	$\frac{35}{384}$	0	$-\frac{500}{1113}$	$\frac{125}{192}$	$-\frac{2187}{6784}$	$\frac{11}{84}$	
u_{l+1}	$\frac{35}{384}$	0	$\frac{500}{1113}$	$\frac{125}{192}$	$-\frac{2187}{6784}$	$\frac{11}{84}$	0
\hat{u}_{l+1}	$\frac{5179}{57600}$	0	$\frac{7571}{16695}$	$\frac{393}{640}$	$-\frac{92097}{339200}$	$\frac{187}{2100}$	$\frac{1}{40}$

который имеет основной метод пятого порядка, а оценщик погрешности — четвертого порядка. Этот метод относится к числу непрерывных, так как позволяет вычислить

$$u(t_l + r\Delta) = u_l + \Delta \sum_{i=1}^6 \sigma_i(r) h_i, \quad 0 \leq r \leq 1,$$

где коэффициенты определяются по формулам

$$\sigma_1(r) = r(1 + r(-1337/480 + r(1039/360 + r(-1163/1152)))),$$

$$\sigma_2(r) = 0,$$

$$\sigma_3(r) = 100r^2(1054/9275 + r(-4682/27825 + r(379/5565)))/3,$$

$$\sigma_4(r) = -5r^2(27/40 + r(-9/5 + r(83/96)))/2,$$

$$\sigma_5(r) = 18225r^2(-3/250 + r(22/375 + r(-37/600)))/848,$$

$$\sigma_6(r) = -22r^2(-3/10 + r(29/30 + r(-17/24)))/7.$$

Таким образом, производится интерполяция пятого порядка и соответственно экстраполяция, что позволяет численно решать не только ОДУ, но и ФДУ.

6.3. Тестовые задачи

Рассмотрим несколько простых примеров систем, относящихся к различным типам ФДУ. В ряде примеров аналитически выписывается точное решение, поэтому численные методы можно тестировать, сравнивая приближенное и точное решение, и по величине погрешности (а также по затратам времени и памяти) делать выводы об эффективности различных численных методов, способов интерполяции и экстраполяции предыстории и способов выбора шага. В некоторых (модельных) примерах аналитическое представление точного решения неизвестно, однако известны полученные аналитическими способами некоторые качественные характеристики поведения решений: асимптотика, предельные циклы. Проведенные численные расчеты позволяют дополнить представления о качественных свойствах решений этих и других моделей. Расчеты проведены методом РКФ 4(5) с автоматическим выбором шага, с интерполяцией вырожденными сплайнами четвертого порядка и с экстраполяцией продолжением. Этот алгоритм наилучшим образом зарекомендовал себя во всех задачах, кроме сильно жестких, и, поэтому, был положен в основу пакета программ Time-Delay System Toolbox [45].

Тест 1.

$$\dot{x}(t) = x(t - \tau),$$

где $t \geq 0$, $\tau = 1$, x - одномерный фазовый вектор с начальными условиями $x(s) = 1$ при $-1 \leq s \leq 0$.

Решение этого уравнения было получено методом шагов в разделе 2.1.2 (см. пример 1.1).

На рис. 2 приведены графики точного (сплошная линия) и приближенного решений (кружки, соответствующие разбиению). Эти обозначения применяются и на всех других рисунках соответствующих тестов.

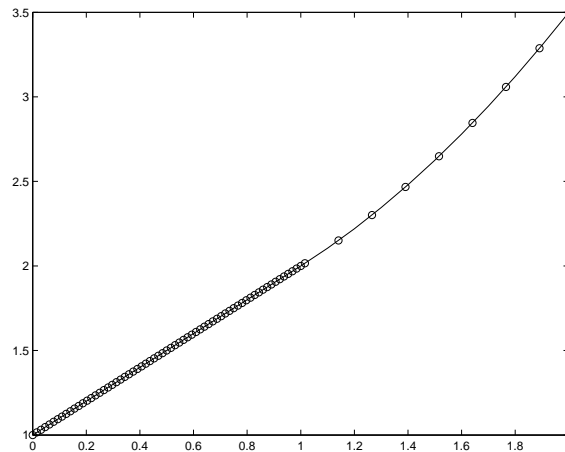


Рис. 2

Тест 2. Рассмотрим уравнение с переменным запаздыванием

$$\dot{x}(t) = x(t - (e^{-t} + 1)) + \cos t - \sin(t - e^{-t} - 1),$$

(см. пример 1.2).

На рис. 3 приведены графики точного и приближенного решений.

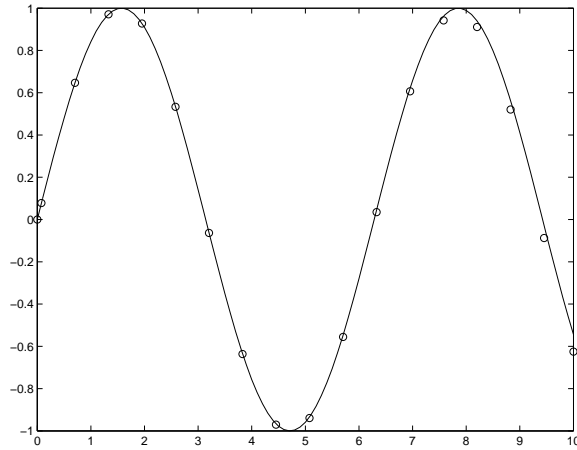


Рис. 3

Тест 3. Рассмотрим уравнение с распределенным запаздыванием

$$\dot{x}(t) = \int_{-1}^0 x_t(s) ds,$$

(см. пример 1.3).

На рис. 4 приведены графики точного и приближенного решений.

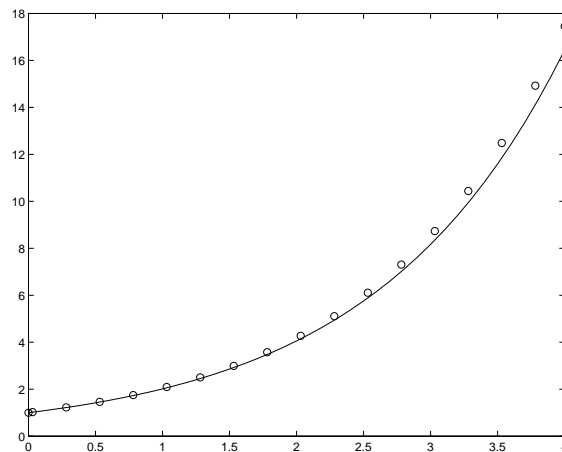


Рис. 4

Тест 4. Линейная система второго порядка, содержащая переменное сосредоточенное и распределённое запаздывания

$$\left\{ \begin{array}{l} \dot{x}_1(t) = -\sin(t)x_1(t) + x_1(t - \frac{t}{2}) - \int_{-\frac{t}{2}}^0 \sin(t+s)x_1(t+s)ds - \\ \quad - e^{\cos(t)}, \\ \dot{x}_2(t) = \cos(t)x_2(t) + x_2(t - \frac{t}{2}) + \int_{-\frac{t}{2}}^0 \cos(t+s)x_2(t+s)ds - \\ \quad - e^{\sin(t)} \end{array} \right.$$

с начальными условиями

$$x_1(t) = e^{\cos(t)}, \quad x_2(t) = e^{\sin(t)} \quad \text{при } t \leq 0.$$

Система рассматривается на отрезке $[0, 2\pi]$ и имеет точное решение, задаваемое теми же формулами, что и начальные условия. На рис. 5 приведены графики двух координат точного и приближенного решений.

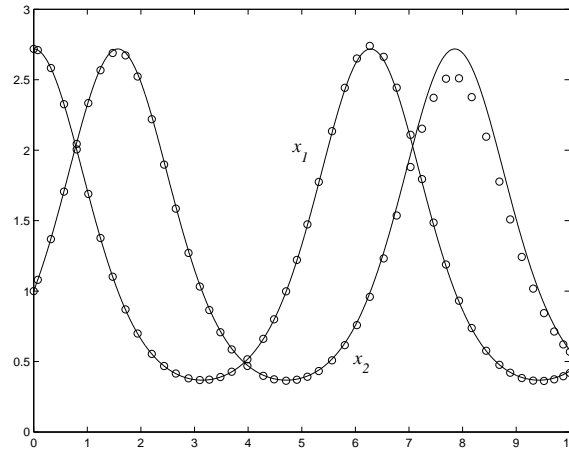


Рис. 5

Тест 5. Нелинейная система второго порядка с распределённым запаздыванием

$$\begin{cases} \dot{x}_1(t) = -\frac{1}{2} \int_{-\pi}^0 x_1(t+s)ds + \frac{2x_1(t) - \frac{\pi}{2}x_2(t)}{\sqrt{x_1^2(t) + x_2^2(t)}}, \\ \dot{x}_2(t) = -\frac{1}{2} \int_{-\pi}^0 x_2(t+s)ds + \frac{2x_2(t) + \frac{\pi}{2}x_1(t)}{\sqrt{x_1^2(t) + x_2^2(t)}} \end{cases}$$

с начальными условиями

$$x_1(t) = t \cos(t), \quad x_2(t) = t \sin(t) \quad \text{при } t \leq 1.$$

Система рассматривается на отрезке $[1, 20]$ и имеет точное решение, задаваемое теми же формулами, что и начальные условия. На рис. 6 приведены графики двух координат точного и приближенного решений, а на рис. 7 — соответствующие траектории в фазовом пространстве.

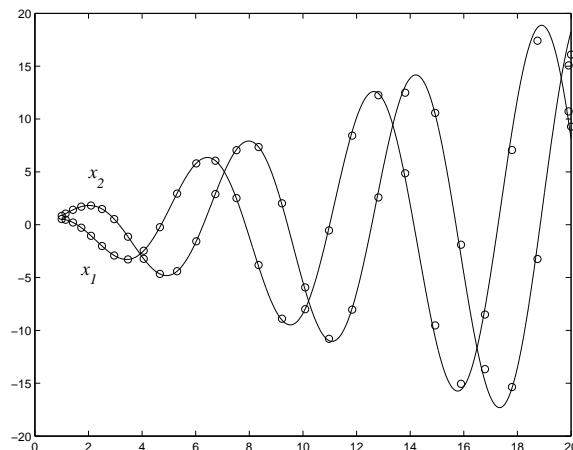


Рис. 6

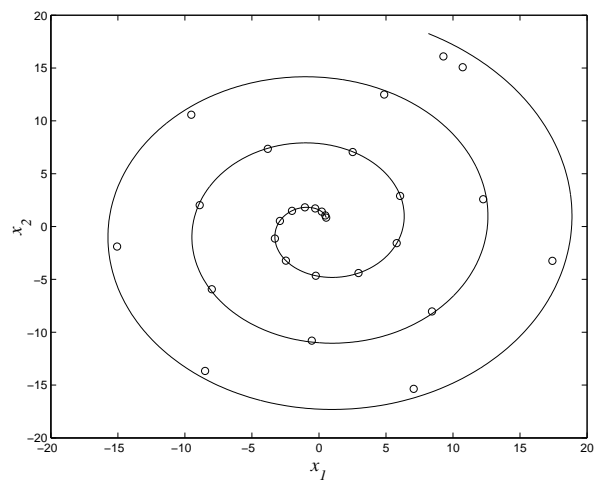


Рис. 7

Тест 6. Линейная система третьего порядка, содержащая постоянное запаздывание

$$\begin{cases} \dot{x}_1(t) = \frac{2}{\pi}(x_1(t) + x_2(t - \frac{\pi}{2})) - x_1(t - \frac{\pi}{2}) - \frac{\pi}{2} \frac{x_2(t)}{x_3(t)}, \\ \dot{x}_2(t) = \frac{2}{\pi}(x_2(t) - x_1(t - \frac{\pi}{2})) - x_2(t - \frac{\pi}{2}) + \frac{\pi}{2} \frac{x_1(t)}{x_3(t)}, \\ \dot{x}_3(t) = \frac{\sqrt{x_1^2(t - \frac{\pi}{2}) + x_2^2(t - \frac{\pi}{2})}}{x_3(t - \frac{\pi}{2})} \end{cases}$$

с начальными условиями

$$x_1(t) = t \cos(t), \quad x_2(t) = t \sin(t), \quad x_3 = t \quad \text{при } t \leq \pi.$$

Система рассматривается на отрезке $[\pi, 20]$ и имеет точное решение, задаваемое теми же формулами, что и начальные условия.

На рис. 8 приведены графики трех координат точного и приближенного решений, на рис. 9 — соответствующие траектории в фазовом пространстве.

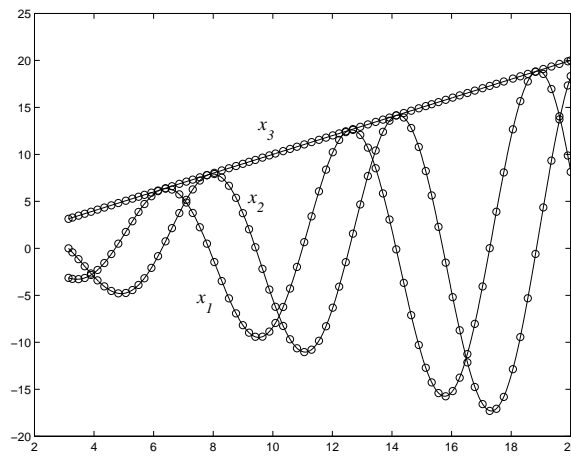


Рис. 8

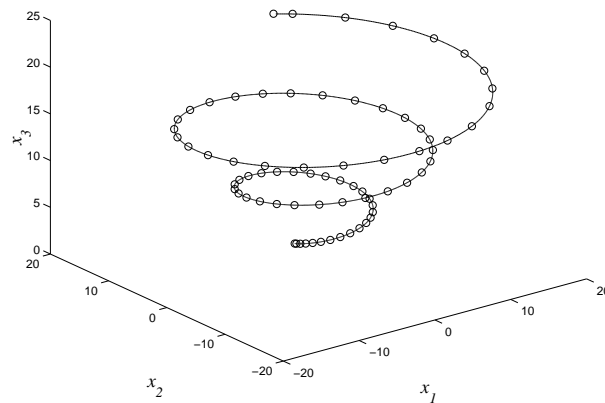


Рис. 9

Тест 7. Жесткая система

$$\begin{cases} \dot{x}_1(t) = \lambda_1 x_1(t) + x_2(t) + x_1(t-1) - \frac{e^{\lambda_2(t-1)}}{\lambda_2 - \lambda_1} - e^{\lambda_1(t-1)}, \\ \dot{x}_2(t) = \lambda_2 x_2(t) \end{cases}$$

с начальными условиями

$$x_1(t) = \frac{e^{\lambda_2 t}}{\lambda_2 - \lambda_1} + e^{\lambda_1 t}, \quad x_2(t) = e^{\lambda_2 t} \quad \text{при } t \leq 0.$$

Система ($\lambda_1 = -100$, $\lambda_2 = -1$) имеет точное решение, задаваемое теми же формулами, что и начальные условия.

Жесткость решений системы ФДУ понимается в следующем (конечномерном) смысле.

Пусть $x(t)$ — решение задачи (6.1) – (6.3). Составим матрицу Якоби функционала f вдоль решения

$$A(t) = \frac{\partial f(t, x(t), x_t(\cdot))}{\partial x}$$

и рассмотрим систему однородных линейных дифференциальных уравнений

$$\dot{x}(t) = A(t)x(t), \quad t \in [t_0, t_0 + \theta].$$

Обозначим через $\lambda_i(t)$, $i = 1, \dots, n$ собственные числа матрицы $A(t)$.

Определение 6.5. Задачу (6.1) – (6.3) назовем жесткой, если

- 1) $\operatorname{Re} \lambda_i(t) < 0$ для всех $t \in [t_0, t_0 + \theta]$, $i = 1, \dots, n$;
- 2) число

$$\sup_{t \in [t_0, t_0 + \theta]} \left\{ \max_{1 \leq i \leq n} |\operatorname{Re} \lambda_i(t)| / \min_{1 \leq i \leq n} |\operatorname{Re} \lambda_i(t)| \right\}$$

велико.

На рис. 10 приведены графики двух координат точного и приближенного решений.

Следует отметить, что интегрирование жестких систем явными методами даже с автоматическим выбором шага приводит к худшим результатам, чем при применении неявных методов (подробнее см. работы [10, ?]). Для сравнения на рис. 11 приведены результаты расчетов, сделанные неявным методом с постоянным шагом.

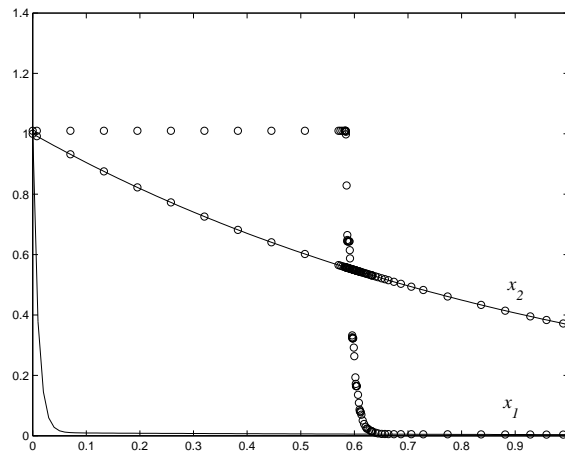


Рис. 10

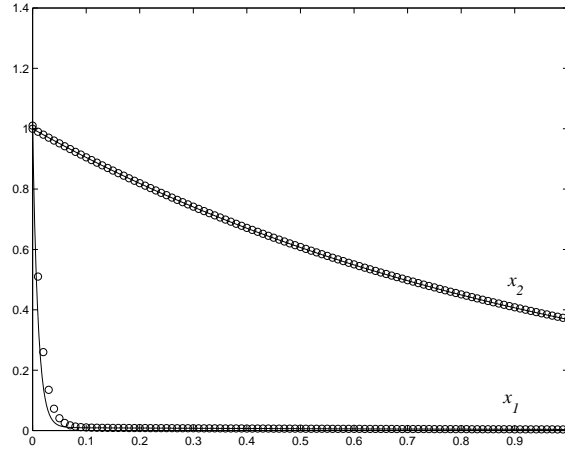


Рис. 11

Тест 8. Система Лотке–Вольтерра с распределённым запаздыванием: [7]

$$\begin{cases} \dot{x}_1(t) = [\varepsilon - \gamma_1 x_2(t) - \int_{-\tau}^0 F_1(s) x_2(t+s) ds] x_1(t) \\ \dot{x}_2(t) = [-\varepsilon + \gamma_2 x_1(t) + \int_{-\tau}^0 F_2(s) x_1(t+s) ds] x_2(t). \end{cases}$$

На рис. 12 приведены графики двух координат приближенного решения, на рис. 13 — соответствующая приближенная траектория в фазовом пространстве. Параметры были взяты следующие: $\tau = 1$, $\varepsilon = 2$, $\gamma_1 = 1$, $\gamma_2 = 1$, $F_1(s) = \sin s$, $F_2(s) = \sin s$, начальные условия $x_1(s) \equiv 1$, $x_2(s) \equiv 1$ при $-\tau \leq s \leq 0$.

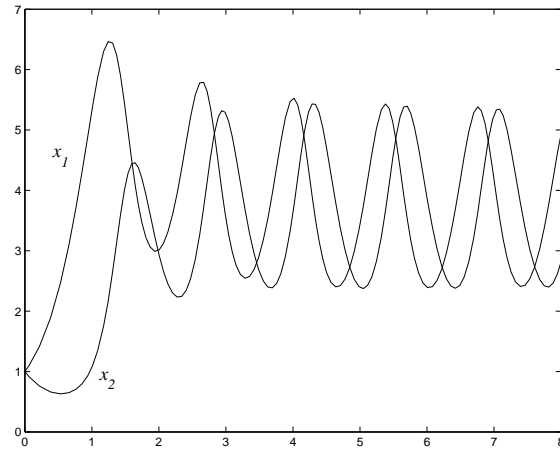


Рис. 12

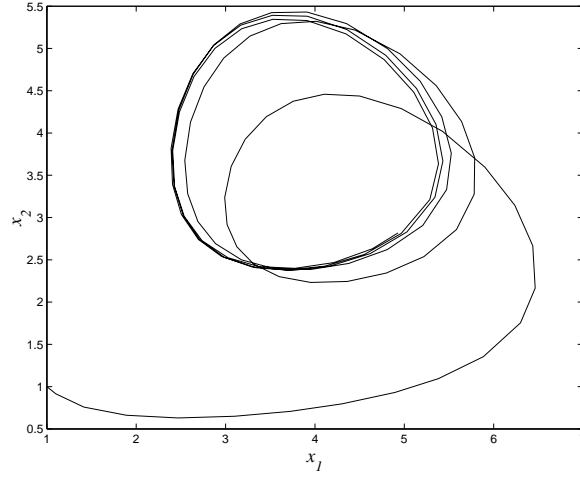


Рис. 13

Тест 9. Уравнение с двумя запаздываниями

$$\dot{N}(t) = N(t)[\gamma - a_0 N(t) - a_1 N(t - \tau_1) + b N(t - \tau_2)],$$

где $\gamma > 0$, $a_0 \geq 0$, $a_1 > 0$, $b > 0$, неотрицательные величины запаздываний τ_1, τ_2 , вообще говоря, несоизмеримы. Пример сообщен I. Györi, которым установлено, что

1) если $a_0 + a_1 > b$ и $a_1 \tau_1$ достаточно мало, то любое положительное решение $N(t)$ стремится к постоянной при $t \rightarrow \infty$;

2) если $a_1 > b$ и $\tau_1 > \tau_2$, то существуют примеры решений $N(t) \rightarrow \infty$ при $t \rightarrow \infty$.

Открытым остается вопрос: если $a_0 + a_1 > b$ и $\tau_1 < \tau_2$, то существуют ли примеры решений $N(t) \rightarrow \infty$ при $t \rightarrow \infty$?

На рис. 14 и 15 приведены графики приближенного решения, соответствующие этим двум случаям. Параметры были взяты следующие: в случае 1) $\gamma = 1$, $a_0 = 1$, $a_1 = 2$, $b = 1$, $\tau_1 = 1$, $\tau_2 = 1$; в случае 2) $\gamma = 0.8$, $a_0 = 0$, $a_1 = 2$, $b = 1$, $\tau_1 = \ln 10$, $\tau_2 = \ln 2$; начальные условия $N(s) \equiv 1$ при $-\tau_1 \leq s \leq 0$.

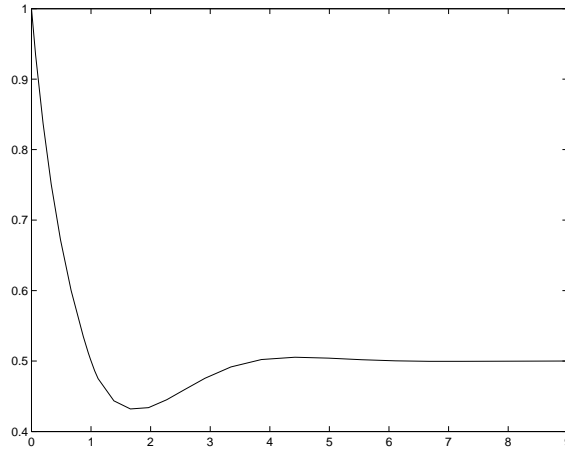


Рис. 14

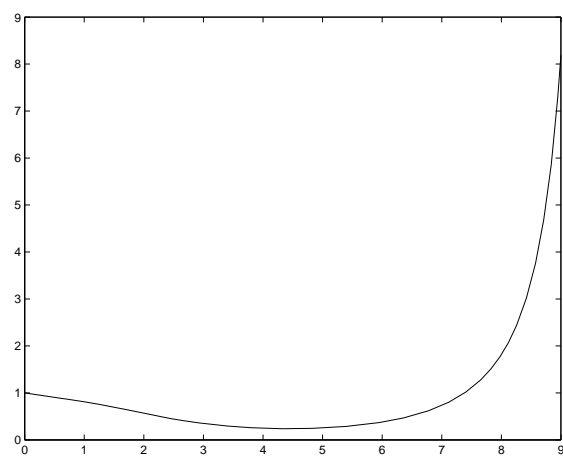


Рис. 15

Список литературы

1. *Азбелев Н.В., Максимов В.П., Разматуллина Л.Ф.* Введение в теорию функционально-дифференциальных уравнений. М. Наука. 1991. 280 с.
2. *Андреева Е.А., Колмановский В.Б., Шайхет Л.Е.* Управление системами с последействием. М. Наука. 1992. 336 с.
3. *Бабский В.Г. и Мышкис А.Д.* Математические модели в биологии, связанные с учетом последействия // В кн. Марри Дж. Нелинейные дифференциальные уравнения в биологии. М. Мир. 1983. С. 383 – 394.
4. *Бейли Н.* Математика в биологии и медицине. М. Мир. 1970. 326 с.
5. *Беллман Р., Кук К.* Дифференциально-разностные уравнения. М. Мир. 1967. 254 с.
6. Бочаров Г.А., Марчук Г.И. Прикладные проблемы математического моделирования в иммунологии // ЖВМ. 2000. Т. 40. N 12. С. 1905 – 1920.
7. *Вольтерра В.* Математическая теория борьбы за существование. М. Наука. 1976. 286 с.
8. *Дибров Б.Ф., Лифшиц М.А., Волькенштейн М.В.* Математическая модель иммунной реакции 1. // Биофизика. 1976. Т. 21. №5. с.905 – 909.
9. *Зверкина Т.С.* Численное интегрирование уравнений с распределенным запаздыванием // Тр. семинара по теории дифференц. уравнений с отклоняющимся аргументом. 1975. Т. 9. С. 82 – 86.
10. *Квон О.Б., Пименов В.Г.* Неявные методы типа Рунге-Кутты для функционально-дифференциальных уравнений // Изв. УрГУ. 1998. №10. С. 69 – 79.
11. *Ким А.В.* i -Гладкий анализ и функционально-дифференциальные уравнения. Екатеринбург. ИММ УрО РАН. 1996. 236 с.
12. *Ким А.В., Пименов В.Г.* О применении i -гладкого анализа к разработке численных методов решения функционально-дифференциальных уравнений // Тр. Ин-та математики и механики УрО РАН. 1998. Т. 5. С. 104 – 126.
13. *Ким А.В., Пименов В.Г.* i -гладкий анализ и численные методы решения функционально-дифференциальных уравнений. Москва-Ижевск. РХД. 2004. 256.
14. *Красовский Н.Н.* Некоторые задачи теории устойчивости движения. М. Гостехиздат. 1959. 211 с.
15. *Марчук Г.И.* Математические модели в иммунологии. М. Наука. 1980. 264 с.
16. *Мышкис А.Д.* Линейные дифференциальные уравнения с запаздывающим аргументом. М. Наука. 1972. 352 с.

17. *Пименов В.Г.* Функционально-дифференциальные уравнения: численные методы. Екатеринбург. Из-во Урал. ун-та. 1998. 80 с.
18. *Пименов В.Г.* Общие линейные методы численного решения дифференциально-функциональных уравнений // Дифференц. уравнения. 2001. Т. 37. №1. С. bibitemRSCb *Романовский Ю.М., Степанова Н.В., Чернавский Д.С.* Математическое моделирование в биофизике. М. Наука. 1975.
19. *Самарский А.А., Гулин А.В.* Численные методы. М. Наука. 1989. 432 с.
20. *Смит Дж.М.* Математические идеи в биологии. М. Мир. 1970. 180 с.
21. *Смит Дж.М.* Модели в экологии. М. Мир. 1976. 184 с.
22. *Хайрер Э., Нерсетт С., Ваннер Г.* Решение обыкновенных дифференциальных уравнений. Нежесткие задачи. М. Мир. 1990. 512 с.
23. *Хайрер Э., Ваннер Г.* Решение обыкновенных дифференциальных уравнений. Т. 2. Жесткие и дифференциально-алгебраические задачи. М. Мир. 1999. 684 с.
24. *Хейл Дж.* Теория функционально-дифференциальных уравнений. М., 1984. 421 с.
25. *Холл Д., Уатт Д.* Современные численные методы решения обыкновенных дифференциальных уравнений. М. Мир. 1979. 312 с.
26. *Эльсгольц Л.Э., Норкин С.Б.* Введение в теорию дифференциальных уравнений с отклоняющимся аргументом. М. Наука. 1971. 296 с.
27. *Baker C.T.H., Bocharov G.A., Filiz A., Ford N.J., Paul C.A.H., Rihan F.A., Tang A., Thomas R.M., Tian H. and Wille D.R.* Numerical modelling by retarded functional differential equations // MCCM tech. rep. №335, University of Manchester. 1998. 35 p.
28. *Baker C.T.H., Bocharov G.A., Paul C.A.H., Rihan F.A.* Modeling and analysis of time-lags in cell proliferations // MCCM tech. rep. №313, University of Manchester. 1997. 16 p.
29. *Baker C.T.H., Bocharov G.A., Rihan F.A.* A report on use of delay differential equations in numerical modeling in the bioscience // MCCM tech. rep. №343, University of Manchester. 1999. 39 p.
30. *Baker C.T.H., Paul C.A.H. and Wille D.R.* Issues in the numerical solution of evolutionary delay differential equations // Advances in Comput. Math. 1995. V. 3. P. 171 – 196.
31. *Baker C.T.H., Paul C.A.H. and Wille D.R.* A bibliography on the numerical solution of delay differential equations // MCCM tech. rep. N 269, University of Manchester. 1995. 52 p.
32. *Bellen A.* Constrained mesh methods for functional differential equations // International Series of Numerical Mathematics, Verlag, Basel. 1985. P. 52 – 70.

33. *Belykh L.N.* On the computational methods in disease models // Mathematical modeling in immunology and medicine (G.I. Marchuk and L.N. Belykh, Eds). North-Holland Publishing company. 1983. P. 79 – 84.
34. *Bocharov G.A.* Modelling the dynamics of LCMV infection in mice: conventional and exhaustive CTV responses // J. Theor. Biol. 1998. V. 192. P. 283 – 304.
35. *Bocharov G.A., Marchuk G.I. Romanyukha A.A.* Numerical solution by LMMs of stiff delay differential systems modelling an immune response // Numer. Math. 1996. V. 73. P. 131 – 146.
36. *Cooke K., Van den Driessche P and Zou X.* Interaction of maturation delay and nonlinear birth in population and epidemic models // J. Math. Biol. V. 39. 1999. 332–252.
37. *Cryer C., Tavernini L.* The numerical solution of Volterra functional differential equations by Euler's method // SIAM J. Numer. Anal. 1972. V.9. P. 105 – 129.
38. *Gopalsamy K.* Stability and oscillations in delay differential equations of population dynamics. The Netherlands. Kluwer Academic Pub. Dordrecht. 1992. 512 p.
39. *Hutchinson G.E.* Circular causal systems in ecology // Ann. N.Y. Acad. Sci. 1948. P. 221–246.
40. *Hairer E., Lubich Ch.* Asymptotic Expansion of the Global Error of Fixed-Stepsize Methods // Numer. Math. 1984. V. 45, N 3. P. 345 – 360.
41. *Iserles A., Norsett S.P.* On the theory of parallel Runge–Kutta methods // IMA J. Numer. Anal. 1990. V. 10. P. 463 – 488.
42. *Kim A.V., Pimenov V.G.* Multistep numerical methods for functional differential equations // Mathematics and Computers in Simulation. 1998. V.45. P. 377 – 384.
43. *Kim A.V., Pimenov V.G.* Numerical Methods for Delay Differential Equations // Lecture Notes Series №44. Seoul National University. Seoul. Korea. 1999. 96 p.
44. *Kolmanovskii V.B., Myshkis A.D.* Applied theory of functional differential equations. Dordrecht – Boston – London. Kluwer Academic Pub. 1992. 236 p.
45. *Kwon W.H., Kim A.V., Pimenov V.G., Lozhnikov A.B., Han S.H., Onegova O.V.* Time-Delay System Toolbox (for use with MATLAB). Beta Version. Seoul National University. Seoul. Korea. 1998. 114 p.
46. *Kuang Y.* Delay differential equations with applications in population dynamics // Mathematics in Science and Engineering. 191. Academic Press. Boston. 1993.
47. *May R.M.* Stability complexity in model ecosystems. Princeton. Princeton Univ. Press. 1973.
48. *MacDonald N.* Time lags in biology models. In Lect. Notes Biomath. Springer. 1978. 112 p.
49. *Ottesen J.T.* Modeling of the Baroflex-Feedback Mechanism with Time-Delay // J. Math. Biol. 1997. V. 36. P. 41 – 63.

50. *Pimenov V.* Asymptotic behavior of global error of general numerical methods for functional differential equations // Stability and Control: Theory and Applications. 2000. Vol. 3. №2. P. 117 – 124.
51. *Shampine L.F.* Conservation Laws and Numerical Solution of ODEs // Comp. and Math. with Appl. 1999. V. 38. P. 61 – 72.
52. *Shampine L.F., Reichelt M.W.* The MATLAB ODE Suite // SIAM Journal on Scientific Computing. 1997. V. 18. P. 21 – 41.
53. *Skeel R.D.* Analysis of Fixed-Stepsize Methods // SIAM J. Numer. Anal. 1976. V. 13. P. 664 – 683.
54. *Tavernini L.* One-step methods for the numerical solution of Volterra functional differential equations // SIAM J. Numer. Anal. 1971. V. 8. P. 786 – 795.