Задача Коши. Методы Рунге-Кутты. Жесткие задачи

Скалько Юрий Иванович
Цыбулин Иван

Задача Коши

Дано обыкновенное дифференциальное уравнение 1го порядка и начальное условие

$$\frac{d\mathbf{y}(t)}{dt} = \mathbf{G}(t, \mathbf{y}(t))$$
$$\mathbf{y}(0) = \mathbf{y}_0$$

Требуется найти решение $\mathbf{y}(t)$ при $t \in [0, T]$

Методы Рунге-Кутты

Методы Рунге-Кутты относятся к *одношаговым методам*, то есть они позволяют по значению решения $\mathbf{u}_n$ вычислить значение в следующей точке $\mathbf{u}_{n+1}$.
Каждый шаг метода состоит из нескольких *стадий*, на которых вычисляются вспомогательные наклоны $\mathbf{k}$. Вычисление наклонов в специально подобранных промежуточных точках позволяет получить метод с высоким порядком аппроксимации.

Общая схема методов Рунге-Кутты

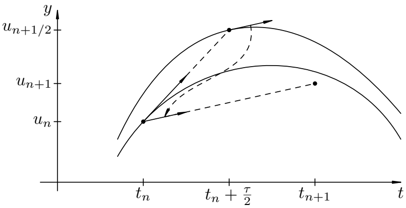
Каждый метод Рунге-Кутты характеризуется набором коэффициентов a_{ij}, b_j, c_i. Один шаг метода проводится по следующей схеме:

$$\mathbf{k}_1 = \mathbf{G}(t_n + c_1\tau, \mathbf{u}_n + \tau\sum_{j=1}^{s} a_{1j}\mathbf{k}_j)$$

$$\vdots$$

$$\mathbf{k}_s = \mathbf{G}(t_n + c_s\tau, \mathbf{u}_n + \tau\sum_{j=1}^{s} a_{sj}\mathbf{k}_j)$$

$$\frac{\mathbf{u}_{n+1} - \mathbf{u}_n}{\tau} = \sum_{j=1}^{s} b_j\mathbf{k}_j$$



$$\mathbf{k}_1 = \mathbf{G}(t_n, \mathbf{u}_n)$$

$$\mathbf{k}_2 = \mathbf{G}\left(t_n + \frac{\tau}{2}, \mathbf{u}_n + \frac{\tau}{2}\mathbf{k}_1\right)$$

$$\frac{\mathbf{u}_{n+1} - \mathbf{u}_n}{\tau} = \mathbf{k}_2$$

Решения, полученные методами Рунге-Кутты

Ниже показаны решения задачи о движении тела в поле тяжести, рассчитанные различными методами Рунге-Кутты с автоматическим выбором длины шага по времени для обеспечения точности $\varepsilon = 10^{-3}$

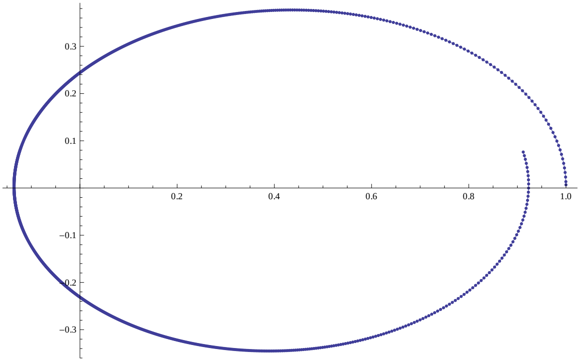


Рис. 1: Метод Эйлера, 1260 шагов

Решения, полученные методами Рунге-Кутты

Ниже показаны решения задачи о движении тела в поле тяжести, рассчитанные различными методами Рунге-Кутты с автоматическим выбором длины шага по времени для обеспечения точности $\varepsilon = 10^{-3}$

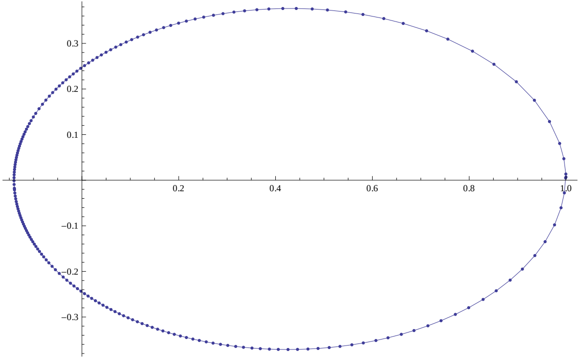


Рис. 1: Явный метод центральной точки, 190($\sim$ 380) шагов

Решения, полученные методами Рунге-Кутты

Ниже показаны решения задачи о движении тела в поле тяжести, рассчитанные различными методами Рунге-Кутты с автоматическим выбором длины шага по времени для обеспечения точности $\varepsilon = 10^{-3}$

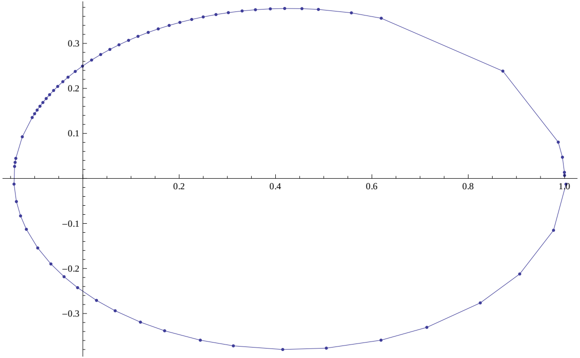


Рис. 1: Метод Рунге-Кутты 4-го порядка, $66(\sim 264)$ шагов

Таблица Бутчера

Коэффициенты a_{ij}, b_j, c_i удобно представлять в виде *таблицы Бутчера*

$$
\begin{array}{c|cccc}
c_1 & a_{11} & a_{12} & \dots & a_{1s} \\
c_2 & a_{21} & a_{22} & \dots & a_{2s} \\
\vdots & & & \ddots & \\
c_s & a_{s1} & a_{s2} & \dots & a_{ss} \\
\hline
 & b_1 & b_2 & \dots & b_s
\end{array}
$$

Таблица Бутчера

Коэффициенты a_{ij}, b_j, c_i удобно представлять в виде *таблицы Бутчера*

$$\begin{array}{c|cccc} c_1 & a_{11} & a_{12} & \dots & a_{1s} \\ c_2 & a_{21} & a_{22} & \dots & a_{2s} \\ \vdots & & & \ddots & \\ c_s & a_{s1} & a_{s2} & \dots & a_{ss} \\ \hline & b_1 & b_2 & \dots & b_s \end{array}$$

Например, явному методу средней точки соответствует таблица

$$\begin{array}{c|cc} 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ \hline & 0 & 1 \end{array}$$

Явные, полуявные и неявные методы Рунге-Кутты

В зависимости от коэффициентов a_{ij} вычисления наклонов $\mathbf{k}_i$ могут происходить по-разному.

- Если матрица a_{ij} имеет ненулевые элементы ниже главной диагонали ($a_{ij} = 0, i \geq j$), то метод называется *явным*. При этом все наклоны $\mathbf{k}_i$ вычисляются через предыдущие без необходимости решать уравнения.
- Если матрица a_{ij} имеет ненулевые элементы и на главной диагонали ($a_{ij} = 0, i > j$), то метод называется *полуявным*. При этом все наклоны $\mathbf{k}_i$ вычисляются последовательно из уравнений.
- Иначе, метод называется *неявным*, и необходимо решать систему уравнений для всех $\mathbf{k}_i$ одновременно.

Разложение наклонов

Поскольку метод Рунге-Кутты определяется своими коэффициентами, можно сформулировать условия на коэффициенты метода, при котором он имеет определенный порядок аппроксимации. Найдем условия первого и второго порядков, для этого подставим $\mathbf{u}_n = [\mathbf{y}]_n$, где $\mathbf{y}(t)$ — решение задачи Коши:

$$
\begin{aligned}
\mathbf{k}_i &= \mathbf{G}\Big(t_n + c_i\tau, [\mathbf{y}]_n + \tau \sum\nolimits_{j=1}^{s} a_{ij}\mathbf{k}_j\Big) = \\
&= [\mathbf{G}]_n + \tau c_i [\mathbf{G}_t]_n + \tau \sum\nolimits_{j=1}^{s} a_{ij}[\mathbf{G}_y]_n \mathbf{k}_j + O(\tau^2) = \\
&= [\mathbf{G}]_n + \tau c_i [\mathbf{G}_t]_n + \tau \sum\nolimits_{j=1}^{s} a_{ij}[\mathbf{G}_y]_n \Big([\mathbf{G}]_n + O(\tau)\Big) + O(\tau^2) = \\
&= [\mathbf{G}]_n + \tau c_i [\mathbf{G}_t]_n + \tau \sum\nolimits_{j=1}^{s} a_{ij}[\mathbf{G}_y]_n [\mathbf{G}]_n + O(\tau^2)
\end{aligned}
$$

Условия первого и второго порядка

Выразим производные $\mathbf{y}'$ и $\mathbf{y}''$ из уравнения: $[\mathbf{y}']_n = [\mathbf{G}]_n, \qquad [\mathbf{y}''] = [\mathbf{G}_t + \mathbf{G}_y\mathbf{G}]_n$

$$\mathbf{k}_i = [\mathbf{G}]_n + \tau c_i [\mathbf{G}_t]_n + \tau \sum_{j=1}^{s} a_{ij} [\mathbf{G}_y]_n [\mathbf{G}]_n + O(\tau^2)$$

$$\sum_{j=1}^{s} b_j k_j = \sum_{j=1}^{s} b_j [\mathbf{G}]_n + \tau \sum_{j=1}^{s} b_j c_j [\mathbf{G}_t]_n + \tau \sum_{i,j=1}^{s} b_i a_{ij} [\mathbf{G}_y]_n [\mathbf{G}]_n$$

$$\frac{[\mathbf{y}]_{n+1} - [\mathbf{y}]_n}{\tau} = [\mathbf{G}]_n + \frac{\tau}{2}[\mathbf{G}_t]_n + \frac{\tau}{2}[\mathbf{G}_y]_n[\mathbf{G}]_n + O(\tau^2)$$

Условия первого и второго порядка

Выразим производные $\mathbf{y}'$ и $\mathbf{y}''$ из уравнения: $[\mathbf{y}']_n = [\mathbf{G}]_n, \qquad [\mathbf{y}''] = [\mathbf{G}_t + \mathbf{G}_y\mathbf{G}]_n$

$$\mathbf{k}_i = [\mathbf{G}]_n + \tau c_i [\mathbf{G}_t]_n + \tau \sum_{j=1}^{s} a_{ij} [\mathbf{G}_y]_n [\mathbf{G}]_n + O(\tau^2)$$

$$\sum_{j=1}^{s} b_j k_j = \sum_{j=1}^{s} b_j [\mathbf{G}]_n + \tau \sum_{j=1}^{s} b_j c_j [\mathbf{G}_t]_n + \tau \sum_{i,j=1}^{s} b_i a_{ij} [\mathbf{G}_y]_n [\mathbf{G}]_n$$

$$\frac{[\mathbf{y}]_{n+1} - [\mathbf{y}]_n}{\tau} = [\mathbf{G}]_n + \frac{\tau}{2}[\mathbf{G}_t]_n + \frac{\tau}{2}[\mathbf{G}_y]_n[\mathbf{G}]_n + O(\tau^2)$$

Условие 1-го порядка аппроксимации: $\displaystyle\sum_{j=1}^{s} b_j = 1.$

Условия 2-го порядка аппроксимации: $\displaystyle\sum_{j=1}^{s} b_j c_j = \sum_{i,j=1}^{s} b_i a_{ij} = \frac{1}{2}.$

Автономизация

Каждой неавтономной задаче Коши

$$\frac{d\mathbf{y}}{dt} = \mathbf{g}(t, \mathbf{y})$$
$$\mathbf{y}\Big|_{t=0} = \mathbf{y}_0$$

можно поставить в соответствие эквивалентную автономную задачу

$$\frac{dt}{ds} = 1$$
$$\frac{d\mathbf{y}}{ds} = \mathbf{g}(t, \mathbf{y})$$
$$\mathbf{t}\Big|_{s=0} = 0$$
$$\mathbf{y}\Big|_{s=0} = \mathbf{y}_0.$$

$$\frac{d\mathbf{Y}}{ds} = \mathbf{G}(\mathbf{Y})$$
$$\mathbf{Y}\Big|_{s=0} = \mathbf{Y}_0$$
$$\mathbf{Y} = (t, \mathbf{y})^\mathsf{T}$$
$$\mathbf{G} = (1, \mathbf{g})^\mathsf{T}.$$

Правило Кутты

Применим метод Рунге-Кутты к автономной системе (заметим, что коэффициенты c_i для автономной системы не используются)

$$\frac{d\mathbf{Y}}{ds} = \mathbf{G}(\mathbf{Y})$$
$$\mathbf{Y}\Big|_{s=0} = \mathbf{Y}_0.$$

Наклоны вычисляются по упрощенным формулам (без явного времени)

$$\mathbf{K}_i = \mathbf{G}(\mathbf{U}_n + \tau \sum_j a_{ij}\mathbf{K}_j) = \begin{pmatrix} 1 \\ \mathbf{g}(\mathbf{U}_n + \tau \sum_j a_{ij}\mathbf{K}_j) \end{pmatrix} \equiv \begin{pmatrix} 1 \\ \mathbf{k}_j \end{pmatrix}$$

Заметим, что первая компонента всех $\mathbf{K}_i$ всегда равна 1.

$$\mathbf{U}_n + \tau \sum_j a_{ij}\mathbf{K}_i = \begin{pmatrix} t_n + \tau \sum_j a_{ij} \\ \mathbf{u}_n + \tau \sum_j a_{ij}\mathbf{k}_j \end{pmatrix}$$

Правило Кутты

В результате применения метода Рунге-Кутты к автономной системе

$$\frac{d\mathbf{Y}}{ds} = \mathbf{G}(\mathbf{Y})$$
$$\mathbf{Y}\Big|_{s=0} = \mathbf{Y}_0.$$

получаются формулы, аналогичные методу Рунге-Кутты для системы

$$\frac{d\mathbf{y}}{dt} = \mathbf{g}(t, \mathbf{y})$$
$$\mathbf{y}\Big|_{t=0} = \mathbf{y}_0.$$

с тем отличием, что вместо коэффициентов c_i в нем стоят величины $\sum_j a_{ij}$. Получается, что если имеется метод Рунге-Кутты, имеющий p-й порядок аппроксимации на автономных системах уравнений, универсальный метод Рунге-Кутты того же порядка получается из него, если положить $c_i = \sum_j a_{ij}$. Последнее условие называется правилом Кутты.

Аппроксимация высших порядков

Намного проще выводить условия порядка для автономных систем, так как в этом случае отсутствуют частные производные по времени.

$$\mathbf{k}_i = \mathbf{G}([\mathbf{y}]_n + \tau \sum_j a_{ij}\mathbf{k}_j) = [\mathbf{G}]_n + \tau[\mathbf{G}_y]_n \sum_j a_{ij}\mathbf{k}_j + \frac{\tau^2}{2}[\mathbf{G}_{yy}]_n \sum_j a_{ij}\mathbf{k}_j \sum_\ell a_{i\ell}\mathbf{k}_\ell + \\ + \frac{\tau^3}{6}[\mathbf{G}_{yyy}]_n \sum_j a_{ij}\mathbf{k}_j \sum_\ell a_{i\ell}\mathbf{k}_\ell \sum_m a_{im}\mathbf{k}_m + O(\tau^4).$$

Запись $[\mathbf{G}_{yyy}]\boldsymbol{\xi}\boldsymbol{\eta}\boldsymbol{\zeta}$ следует понимать как свертку

$$[\mathbf{G}_{yyy}]\boldsymbol{\xi}\boldsymbol{\eta}\boldsymbol{\zeta} = \sum_{j,\ell,m} \frac{\partial^3 G_i}{\partial y_j \partial y_\ell \partial y_m} \xi_j \eta_\ell \zeta_m.$$

В общем случае, за производной порядка q следует q векторов, с которой она сворачивается. Из-за равенства смешанных производных, порядок векторов не важен.

Аппроксимация высших порядков

Из разложения до порядка $O(\tau^4)$

$$\mathbf{k}_i = [\mathbf{G}]_n + \tau \sum_j a_{ij}[\mathbf{G}_y]\mathbf{k}_j + \frac{\tau^2}{2} \sum_{j\ell} a_{ij}a_{i\ell}[\mathbf{G}_{yy}]\mathbf{k}_j\mathbf{k}_\ell + \frac{\tau^3}{6} \sum_{j\ell m} a_{ij}a_{i\ell}a_{im}[\mathbf{G}_{yyy}]\mathbf{k}_j\mathbf{k}_\ell\mathbf{k}_m + O(\tau^4).$$

отбрасыванием степеней можно получить более грубые разложения

$$\mathbf{k}_i = [\mathbf{G}]_n + \tau \sum_j a_{ij}[\mathbf{G}_y]\mathbf{k}_j + \frac{\tau^2}{2} \sum_{j\ell} a_{ij}a_{i\ell}[\mathbf{G}_{yy}]\mathbf{k}_j\mathbf{k}_\ell + O(\tau^3) \tag{1}$$

$$\mathbf{k}_i = [\mathbf{G}]_n + \tau \sum_j a_{ij}[\mathbf{G}_y]\mathbf{k}_j + O(\tau^2) \tag{2}$$

$$\mathbf{k}_i = [\mathbf{G}]_n + O(\tau) \tag{3}$$

Подставим (3) в (2):

$$\mathbf{k}_i = [\mathbf{G}]_n + \tau \textstyle\sum_j a_{ij}[\mathbf{G}_y]_n([\mathbf{G}]_n + O(\tau)) + O(\tau^2) = [\mathbf{G}]_n + \tau \textstyle\sum_j a_{ij}[\mathbf{G}_y\mathbf{G}]_n + O(\tau^2)$$

Аппроксимация высших порядков

Постепенно избавляемся от $\mathbf{k}_i$ в правой части разложений

$$\mathbf{k}_i = [\mathbf{G}]_n + O(\tau) \tag{4}$$

$$\mathbf{k}_i = [\mathbf{G}]_n + \tau \sum_j a_{ij} [\mathbf{G}_y \mathbf{G}]_n + O(\tau^2) \tag{5}$$

$$\mathbf{k}_i = [\mathbf{G}]_n + \tau \sum_j a_{ij} [\mathbf{G}_y]_n \mathbf{k}_j + \frac{\tau^2}{2} \sum_{j\ell} a_{ij} a_{i\ell} [\mathbf{G}_{yy}]_n \mathbf{k}_j \mathbf{k}_\ell + O(\tau^3) \tag{6}$$

В произведении $[\mathbf{G}_y y]_n \mathbf{k}_j \mathbf{k}_\ell$ из (6) величины $\mathbf{k}_j$ и $\mathbf{k}_\ell$ могут быть взяты с погрешностью $O(\tau)$, так как умножаются на τ^2, а величина $[\mathbf{G}_y]_n \mathbf{k}_j$ требует значения $\mathbf{k}_j$ с точностью до $O(\tau^2)$ так как умножается всего лишь на первую степень τ.

Аппроксимация высших порядков

Постепенно избавляемся от $\mathbf{k}_i$ в правой части разложений. Величины вида $\sum_j a_{ij}$ сразу для краткости заменим на c_i:

$$\mathbf{k}_i = [\mathbf{G}]_n + O(\tau)$$
$$\mathbf{k}_i = [\mathbf{G}]_n + \tau \sum_j a_{ij} [\mathbf{G}_y \mathbf{G}]_n + O(\tau^2) = [\mathbf{G}]_n + \tau c_i [\mathbf{G}_y \mathbf{G}]_n + O(\tau^2)$$

$$\mathbf{k}_i = [\mathbf{G}]_n + \tau \sum_j a_{ij} [\mathbf{G}_y]_n \mathbf{k}_j + \frac{\tau^2}{2} \sum_{j\ell} a_{ij} a_{i\ell} [\mathbf{G}_{yy}]_n \mathbf{k}_j \mathbf{k}_\ell + O(\tau^3) = [\mathbf{G}]_n +$$
$$\tau \sum_j a_{ij} [\mathbf{G}_y]_n ([\mathbf{G}]_n + \tau c_i [\mathbf{G}_y \mathbf{G}]_n + O(\tau^2)) + \frac{\tau^2}{2} \sum_{j\ell} a_{ij} a_{i\ell} [\mathbf{G}_{yy}]_n ([\mathbf{G}]_n + O(\tau))([\mathbf{G}]_n + O(\tau)) + O(\tau^3) =$$
$$= [\mathbf{G}]_n + c_i \tau [\mathbf{G}_y \mathbf{G}]_n + \tau^2 \sum_j a_{ij} c_j [\mathbf{G}_y \mathbf{G}_y \mathbf{G}]_n + \frac{\tau^2 c_i^2}{2} [\mathbf{G}_{yy} \mathbf{G}\mathbf{G}]_n + O(\tau^3)$$

Аппроксимация высших порядков

На данный момент получены разложения $\mathbf{k}_i$ до третьего порядка:

$$\mathbf{k}_i = [\mathbf{G}]_n + O(\tau)$$
$$\mathbf{k}_i = [\mathbf{G}]_n + \tau c_i [\mathbf{G}_y \mathbf{G}]_n + O(\tau^2)$$
$$\mathbf{k}_i = [\mathbf{G}]_n + \tau c_i [\mathbf{G}_y \mathbf{G}]_n + \tau^2 \sum_j a_{ij} c_j [\mathbf{G}_y \mathbf{G}_y \mathbf{G}]_n + \frac{\tau^2 c_i^2}{2} [\mathbf{G}_{yy} \mathbf{G}\mathbf{G}]_n + O(\tau^3)$$

Проделывая аналогичные подстановки в разложения четвертого порядка, получаем

$$\mathbf{k}_i = [\mathbf{G}]_n + \tau c_i [\mathbf{G}_y \mathbf{G}]_n + \tau^2 \sum_j a_{ij} c_j [\mathbf{G}_y \mathbf{G}_y \mathbf{G}]_n + \frac{\tau^2 c_i^2}{2} [\mathbf{G}_{yy} \mathbf{G}\mathbf{G}]_n + \tau^3 \sum_{j\ell} a_{ij} a_{j\ell} c_\ell [\mathbf{G}_y \mathbf{G}_y \mathbf{G}_y \mathbf{G}]_n +$$
$$+ \frac{\tau^3}{2} \sum_j a_{ij} c_j^2 [\mathbf{G}_y \mathbf{G}_{yy} \mathbf{G}\mathbf{G}]_n + \tau^3 c_i \sum_j a_{ij} c_j [\mathbf{G}_{yy} \mathbf{G}_y \mathbf{G}\mathbf{G}]_n + \frac{\tau^3}{6} c_i^3 [\mathbf{G}_{yyy} \mathbf{G}\mathbf{G}\mathbf{G}]_n + O(\tau^4).$$

В последнем выражении учтено, что $[\mathbf{G}_{yy}(\mathbf{G}_y \mathbf{G})\mathbf{G}]_n = [\mathbf{G}_{yy}\mathbf{G}(\mathbf{G}_y \mathbf{G})]_n$ из-за симметрии $\mathbf{G}_{yy}$.

Аппроксимация высших порядков

В выражении

$$\mathbf{k}_i = [\mathbf{G}]_n + \tau c_i [\mathbf{G}_y \mathbf{G}]_n + \tau^2 \sum_j a_{ij} c_j [\mathbf{G}_y \mathbf{G}_y \mathbf{G}]_n + \frac{\tau^2 c_i^2}{2} [\mathbf{G}_{yy} \mathbf{G}\mathbf{G}]_n + \tau^3 \sum_{j\ell} a_{ij} a_{j\ell} c_\ell [\mathbf{G}_y \mathbf{G}_y \mathbf{G}_y \mathbf{G}]_n +$$
$$+ \frac{\tau^3}{2} \sum_j a_{ij} c_j^2 [\mathbf{G}_y \mathbf{G}_{yy} \mathbf{G}\mathbf{G}]_n + \tau^3 c_i \sum_j a_{ij} c_j [\mathbf{G}_{yy} \mathbf{G}_y \mathbf{G}\mathbf{G}]_n + \frac{\tau^3}{6} c_i^3 [\mathbf{G}_{yyy} \mathbf{G}\mathbf{G}\mathbf{G}]_n + O(\tau^4).$$

имеются 8 разных произведений, сожержащих **G** и ее производные. Соберем коэффициенты перед ними в таблицу:

	$\mathbf{G}$	$\tau \mathbf{G}_y \mathbf{G}$	$\tau^2 \mathbf{G}_y \mathbf{G}_y \mathbf{G}$	$\tau^2 \mathbf{G}_{yy} \mathbf{G}\mathbf{G}$	$\tau^3 \mathbf{G}_y \mathbf{G}_y \mathbf{G}_y \mathbf{G}$	$\tau^3 \mathbf{G}_y \mathbf{G}_{yy} \mathbf{G}\mathbf{G}$	$\tau^3 \mathbf{G}_{yy} \mathbf{G}_y \mathbf{G}\mathbf{G}$	$\tau^3 \mathbf{G}_{yyy} \mathbf{G}\mathbf{G}\mathbf{G}$
$\sum_i b_i \mathbf{k}_i$	$\sum_i b_i$	$\sum_i b_i c_i$	$\sum_{ij} b_i a_{ij} c_j$	$\sum_i b_i \frac{c_i^2}{2}$	$\sum_{ij\ell} b_i a_{ij} a_{j\ell} c_\ell$	$\sum_j b_i a_{ij} \frac{c_j^2}{2}$	$\sum_{ij} b_i c_i a_{ij} c_j$	$\sum_i b_i \frac{c_i^3}{6}$

Аппроксимация высших порядков

	$\mathbf{G}$	$\tau\mathbf{G}_y\mathbf{G}$	$\tau^2\mathbf{G}_y\mathbf{G}_y\mathbf{G}$	$\tau^2\mathbf{G}_{yy}\mathbf{G}\mathbf{G}$	$\tau^3\mathbf{G}_y\mathbf{G}_y\mathbf{G}_y\mathbf{G}$	$\tau^3\mathbf{G}_y\mathbf{G}_{yy}\mathbf{G}\mathbf{G}$	$\tau^3\mathbf{G}_{yy}\mathbf{G}_y\mathbf{G}\mathbf{G}$	$\tau^3\mathbf{G}_{yyy}\mathbf{G}\mathbf{G}\mathbf{G}$
$\sum_i b_i\mathbf{k}_i$	$\sum_i b_i$	$\sum_i b_i c_i$	$\sum_{ij} b_i a_{ij} c_j$	$\sum_i b_i \frac{c_i^2}{2}$	$\sum_{ij\ell} b_i a_{ij} a_{j\ell} c_\ell$	$\sum_j b_i a_{ij} \frac{c_j^2}{2}$	$\sum_{ij} b_i c_i a_{ij} c_j$	$\sum_i b_i \frac{c_i^3}{6}$

Разложим аналогично $\frac{[\mathbf{y}]_{n+1}-[\mathbf{y}]_n}{\tau}$:

$$\frac{[\mathbf{y}]_{n+1}-[\mathbf{y}]_n}{\tau} = [\mathbf{y}']_n + \frac{\tau}{2}[\mathbf{y}'']_n + \frac{\tau^2}{6}[\mathbf{y}''']_n + \frac{\tau^3}{24}[\mathbf{y}^{IV}]_n + O(\tau^4)$$

и учтем, что

$$\mathbf{y}'(t) = \mathbf{G}(\mathbf{y}(t)) = \mathbf{G}$$

$$\mathbf{y}''(t) = \frac{d}{dt}\mathbf{G}(\mathbf{y}(t)) = \mathbf{G}_y\mathbf{y}'(t) = \mathbf{G}_y\mathbf{G}$$

$$\mathbf{y}'''(t) = (\mathbf{G}_y\mathbf{G})' = \mathbf{G}_{yy}\mathbf{G}\mathbf{G} + \mathbf{G}_y\mathbf{G}_y\mathbf{G}$$

$$\mathbf{y}^{IV}(t) = (\mathbf{G}_{yy}\mathbf{G}\mathbf{G} + \mathbf{G}_y\mathbf{G}_y\mathbf{G})' = \mathbf{G}_{yyy}\mathbf{G}\mathbf{G}\mathbf{G} + 3\mathbf{G}_{yy}\mathbf{G}_y\mathbf{G}\mathbf{G} + \mathbf{G}_y\mathbf{G}_{yy}\mathbf{G}\mathbf{G} + \mathbf{G}_y\mathbf{G}_y\mathbf{G}_y\mathbf{G}$$

Аппроксимация высших порядков

Допишем в таблицу коэффициенты разожения $\frac{[\mathbf{y}]_{n+1}-[\mathbf{y}]_n}{\tau}$

	$\mathbf{G}$	$\tau\mathbf{G}_y\mathbf{G}$	$\tau^2\mathbf{G}_y\mathbf{G}_y\mathbf{G}$	$\tau^2\mathbf{G}_{yy}\mathbf{G}\mathbf{G}$	$\tau^3\mathbf{G}_y\mathbf{G}_y\mathbf{G}_y\mathbf{G}$	$\tau^3\mathbf{G}_y\mathbf{G}_{yy}\mathbf{G}\mathbf{G}$	$\tau^3\mathbf{G}_{yy}\mathbf{G}_y\mathbf{G}\mathbf{G}$	$\tau^3\mathbf{G}_{yyy}\mathbf{G}\mathbf{G}\mathbf{G}$
$\sum_i b_i\mathbf{k}_i$	$\sum_i b_i$	$\sum_i b_i c_i$	$\sum_{ij} b_i a_{ij} c_j$	$\sum_i b_i \frac{c_i^2}{2}$	$\sum_{ij\ell} b_i a_{ij} a_{j\ell} c_\ell$	$\sum_j b_i a_{ij} \frac{c_j^2}{2}$	$\sum_{ij} b_i c_i a_{ij} c_j$	$\sum_i b_i \frac{c_i^3}{6}$
$\frac{[\mathbf{y}]_{n+1}-[\mathbf{y}]_n}{\tau}$	1	$\frac{1}{2}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{24}$	$\frac{1}{24}$	$\frac{3}{24}$	$\frac{1}{24}$

Из этой таблицы можно получить условия аппроксимации вплоть до четвертого порядка. Для этого необходимо приравнять выражения во второй строке со значениями третьей. Для аппроксимации порядка p нужно оставить только столбцы с τ^q где $q < p$.

$$
\begin{array}{ll}
O(\tau) & \sum_i b_i = 1 \\
O(\tau^2) & \sum_i b_i c_i = 1/2 \\
O(\tau^3) & \sum_{ij} b_i a_{ij} c_j = 1/6 \\
O(\tau^3) & \sum_i b_i c_i^2 = 1/3
\end{array}
\qquad
\begin{array}{ll}
O(\tau^4) & \sum_{ij\ell} b_i a_{ij} a_{j\ell} c_\ell = 1/24 \\
O(\tau^4) & \sum_{ij} b_i a_{ij} c_j^2 = 1/12 \\
O(\tau^4) & \sum_{ij} b_i c_i a_{ij} c_j = 1/8 \\
O(\tau^4) & \sum_i b_i c_i^3 = 1/4
\end{array}
$$

Барьеры Бутчера

Бутчер доказал несколько теорем о связи порядка аппроксимации и количества стадий у методов Рунге-Кутты. Явные методы с $s < 5$ стадиями могут иметь порядок не выше s, но после $s = 5$ стадий наступает, так называемый, *первый барьер Бутчера*, и порядок аппроксимации не превышает $s - 1$. При увеличении s возникают все новые барьеры, понижающие порядок аппроксимации.

Однако, для неявных методов ограничение не такое строгое. Например есть семейство методов (Гаусса), у которых порядок аппроксимации $2s$ при любом числе стадий.

Устойчивость

Если правая часть ОДУ $G(t, y)$ липшицева по y с константой L

$$\|\mathbf{G}(t, \mathbf{y}) - \mathbf{G}(t, \mathbf{v})\| \leq L\|\mathbf{y} - \mathbf{v}\|,$$

то несложно показать, что константа устойчивости для методов Рунге-Кутты порядка $C \sim \exp\{O(L)T\}$. Следовательно, имеет место сходимость решения разностной задачи к решению дифференциальной задачи.

Но в случаях, когда $LT \gg 1$ константа устойчивости становится огромной. Вспомним, что ошибка сходимости связана с ошибкой аппроксимации соотношением

$$\varepsilon_{\text{сх}} = C\varepsilon_{\text{аппр}}(\tau).$$

Для того, чтобы обеспечить малую ошибку сходимости необходимо выбирать очень маленький шаг по времени τ.

Жесткие задачи

Жесткие системы ОДУ описывают, как правило, одновременно проходящие очень быстрые и очень медленные процессы. Например, в задачах химической кинетики бывают различия в скоростях реакций до 10^{15} раз.

Жесткие задачи

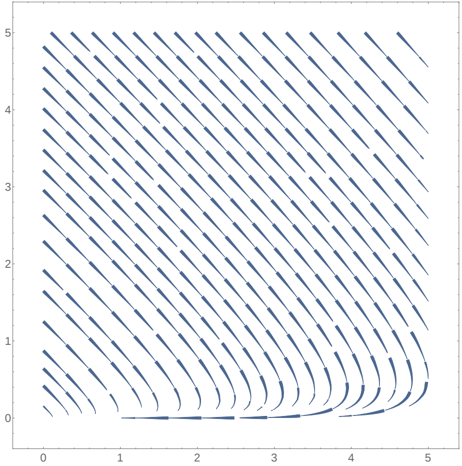
Жесткие системы ОДУ описывают, как правило, одновременно проходящие очень быстрые и очень медленные процессы. Например, в задачах химической кинетики бывают различия в скоростях реакций до 10^{15} раз. Оказывается, что быстро протекающие процессы, даже быстро закончившись, продолжают влиять на численное решение задачи, вынуждая рассчитывать решение с очень малым шагом по времени, где это, казалось бы, совершенно не требуется (решение довольно гладкое).

Поле решений жесткой задачи

Поле решений уравнения

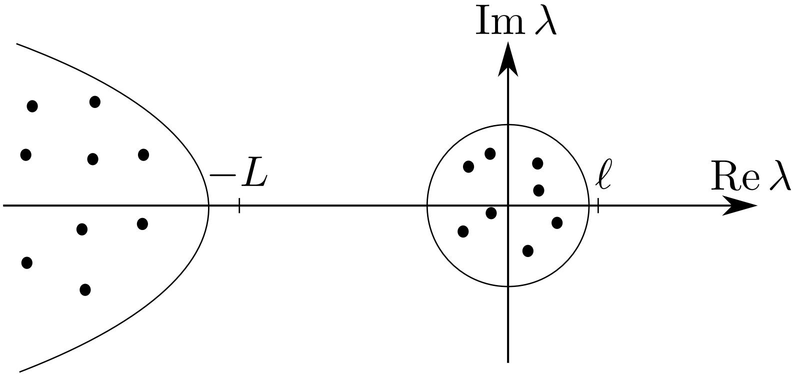
$$\dot{x} = -0.5x + 20y$$
$$\dot{y} = -20y$$

содержит резкие повороты — индикатор жесткости задачи.



Определение жесткой задачи

Можно дать следующее определение:
Жесткая задача — это такая задача, у которой собственные числа матрицы Якоби G_y разбиваются на две части — мягкую часть спектра $|\lambda_i| < \ell$ и жесткую часть спектра $\operatorname{Re}\Lambda_i < -L$, причем $\ell \ll L$. Величина L/ℓ называется показателем жесткости.



Модельное уравнение

Выяснить, как на жестких задачах себя ведет тот или иной метод, можно на модельном уравнении

$$y' = \lambda y, \quad \operatorname{Re}\lambda < 0$$

Все линейные численные методы для решения этого уравнения будут иметь вид

$$u_{n+1} = r(\lambda\tau)u_n,$$

где $r(z)$ — функция, зависящая только от метода. Эта функция называется *функцией устойчивости метода*. Если при данном сочетании λ и τ значение функции $r(\lambda\tau)$ по модулю больше единицы, решение будет экспоненциально возрастать, что противоречит реальному поведению решения при $\operatorname{Re}\lambda < 0$. Область комплексной плоскости $\mathbb{C}$, в которой $|r(z)| < 1$ называется *областью устойчивости метода*.

Функция и область устойчивости

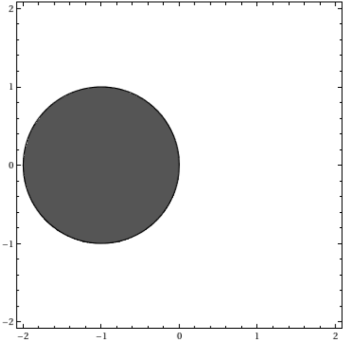
Если при данном τ вся жесткая часть спектра попадает в область устойчивости, она гарантированно не будет экспоненциально возрастать, и решать систему ОДУ можно только обращая внимание на мягкую часть спектра.

Для явного метода Эйлера

$$\frac{u_{n+1} - u_n}{\tau} = \lambda u_n$$

функция устойчивости

$$u_{n+1} = (1 + \tau\lambda)u_n = (1 + z)u_n$$
$$r(z) = 1 + z$$



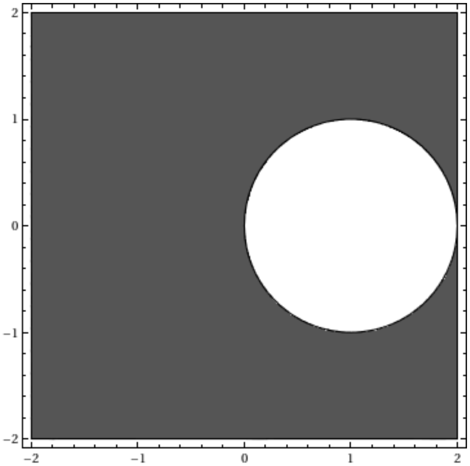
Область устойчивости

Для неявного метода Эйлера

$$\frac{u_{n+1} - u_n}{\tau} = \lambda u_{n+1}$$

функция устойчивости

$$u_{n+1} = \frac{u_n}{1 - \tau\lambda} = \frac{u_n}{1 - z}$$
$$r(z) = \frac{1}{1 - z}$$

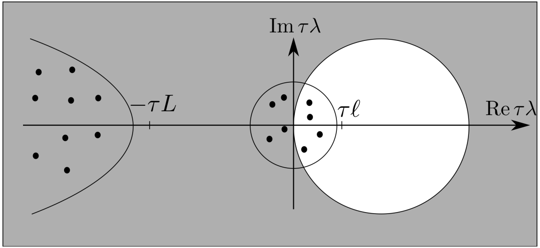


Допустимый шаг τ для жесткой задачи

Если шаг по времени τ таков, что все собственные числа Λ_i задачи из жесткой части спектра попадают в область устойчивости данного метода

$$|r(\tau\Lambda_i)| \leq 1,$$

то с таким шагом решать жесткую задачу можно.
Если это требование нарушить, жесткие компоненты решения начнут экспоненциально возрастать (хотя обязаны стремится к нулю!).



A- и L-устойчивость

По виду области устойчивости методы можно дополнительно классифицировать. Это позволяет выбирать метод, наиболее подходящий для конкретного вида жесткой части спектра задачи.

- A-устойчивость означает, что во всей полуплоскости $\operatorname{Re} z < 0$ метод устойчив, т.е. $|r(z)| < 1$. Такой меотд годится для любых жестких задач.
- $A(\alpha)$-устойчивость означает, что в конусе $|\operatorname{Im} z| < -\operatorname{tg}\alpha \operatorname{Re} z$ метод устойчив. A-устойчивость эквивалентна $A(90^\circ)$. Такой метод годится для задач, у который жесткий спектр прижат к действительной оси. Чем больше α, тем универсальнее метод.
- L-устойчивость означает, что $\lim_{z\to-\infty} r(z) = 0$. Это свойство говорит, что при большом шаге τ жесткая часть спектра стремится к нулю достаточно быстро. Эти методы хороши тем, что допускают интегрирование погранслоя с большим шагом. Не L-устойчивые методы осциллируют при выходе из погранслоя.

A- и L-устойчивые методы

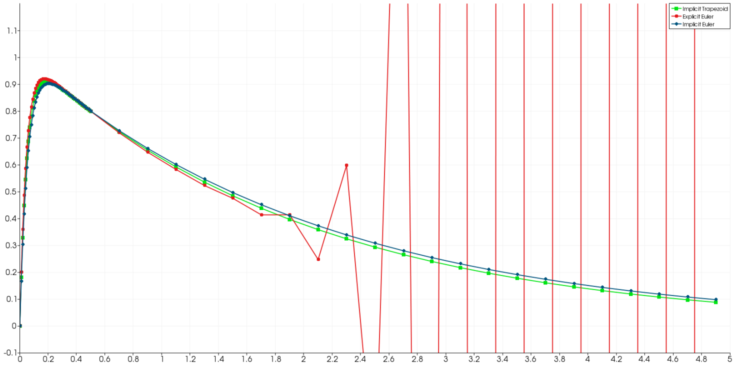


Рис. 3: Решение жесткой задачи разными методами с ограничением шага в погранслое (красный — не A-устойчивый явный метод Эйлера, зеленый — A-, но не L-устойчивый неявный метод средней точки, синий — A- и L-устойчивый неявный метод Эйлера)

A- и L-устойчивые методы

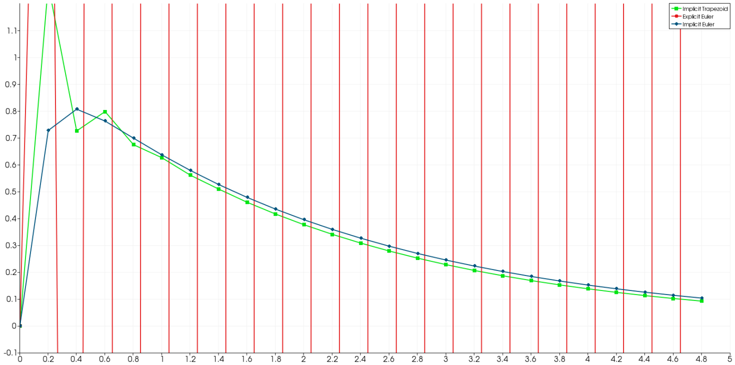


Рис. 3: Решение жесткой задачи разными методами с большим шагом (красный — не A-устойчивый явный метод Эйлера, зеленый — A-, но не L-устойчивый неявный метод средней точки, синий — A- и L-устойчивый неявный метод Эйлера)

Функция устойчивости методов Рунге-Кутты

Для методов Рунге-Кутты функцию устойчивости можно вычислить по формуле

$$r(z) = \frac{\det(\mathbf{E} - z\mathbf{A} + z\mathbf{1}\mathbf{b}^T)}{\det(\mathbf{E} - z\mathbf{A})},$$

где $\mathbf{1}$ — вектор из единиц.

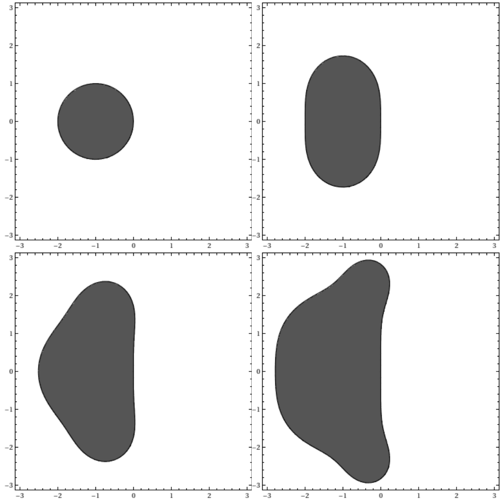
Функция устойчивости методов Рунге-Кутты

Для методов Рунге-Кутты функцию устойчивости можно вычислить по формуле

$$r(z) = \frac{\det(\mathbf{E} - z\mathbf{A} + z\mathbf{1}\mathbf{b}^T)}{\det(\mathbf{E} - z\mathbf{A})},$$

где $\mathbf{1}$ — вектор из единиц. Для случая явного метода, $r(z)$ является многочленом от z степени s (число стадий). Но также, $r(z)$ должен с точностью до $O(z^{p+1})$ совпадать с разложением e^z в ряд по z (аппроксимация порядка p). Если $s = p$, то $r(z)$ есть просто первые $s + 1$ членов ряда Тейлора функции e^z.

Области устойчивости методов Рунге-Кутты 1–4 порядка



Спасибо за внимание!

Цыбулин Иван
e-mail: tsybulin@crec.mipt.ru