АДДИТИВНЫЙ МЕТОД ВТОРОГО ПОРЯДКА ДЛЯ РЕШЕНИЯ ЖЕСТКИХ ЗАДАЧ

Новиков Евгений Александрович
Институт вычислительного моделирования СО РАН
e-mail: novikov@icm.krasn.ru

Для численного решения жестких задач **(1)**

$$y' = f(t, y), \quad y(t_0) = y_0, \quad t_0 \le t \le t_k \qquad \textbf{(1)}$$

обычно применяются ***L***-устойчивые методы. Здесь ***y*** и ***f*** – вещественные ***N***-мерные вектор-функции, ***t*** – независимая переменная.

В случае большой задачи **(1)** для методов с неограниченной областью устойчивости общие вычислительные затраты фактически полностью определяются временем декомпозиции матрицы Якоби системы **(1)**.

Во многих алгоритмах используется замораживание матрицы Якоби, то есть применение одной матрицы на нескольких шагах интегрирования. Это позволяет значительно уменьшить вычислительные затраты.

Наиболее естественно это осуществляется **в итерационных методах решения** обыкновенных дифференциальных уравнений, где данная матрица не влияет на порядок точности численной схемы, а только определяет скорость сходимости итерационного процесса.

Для безытерационных методов вопрос о замораживании или какой-либо другой аппроксимации матрицы Якоби значительно более сложный.

Задачу **(1)** можно записать в виде **(2)**

$$y' = [\,f(t, y) - By] + By,\ \ y(t_0) = y_0,\ t_0 \le t \le t_k, \quad \textbf{(2)}$$

где ***B*** есть некоторая аппроксимация матрицы Якоби.

Предполагая, что вся жесткость сосредоточена в слагаемом ***By***, выражение в квадратных скобках можно интерпретировать как нежесткую часть.

Если при построении безытерационных методов учитывать этот факт, то в алгоритмах интегрирования можно использовать замораживание матрицы Якоби, которая может вычисляться как аналитически, так и численно.

Для некоторых задач в качестве матрицы ***B*** можно использовать симметричную часть матрицы Якоби или применять ее диагональную аппроксимацию.

Здесь построен четырехстадийный метод второго порядка точности, допускающий различные виды аппроксимации матрицы Якоби. Получены оценка ошибки и неравенство для контроля точности вычислений.

ЧИСЛЕННАЯ СХЕМА ДЛЯ АВТОНОМНЫХ ЗАДАЧ

Рассмотрим автономную задачу Коши **(3)**

$$y' = \varphi(y) + g(y), \; y(t_0) = y_0, \; t_0 \le t \le t_k, \quad (3)$$

где y, φ и g – вещественные N-мерные вектор-функции, t – независимая переменная. Будем полагать, что вся жесткость сосредоточена в функции $g(y)$, а $\varphi(y)$ есть нежесткая часть.

Для численного решения **(3)** рассмотрим метод **(4)**

$$y_{n+1} = y_n + \sum_{i=1}^{4} p_i k_i, \; D_n = E - ahg'_n,$$

$$k_1 = h\varphi(y_n), \; Dk_2 = h[\varphi(y_n) + g(y_n)], \quad (4)$$

$$Dk_3 = k_2, \; k_4 = h\varphi(y_n + \beta_{41}k_1 + \beta_{42}k_2 + \beta_{43}k_3),$$

где E – единичная матрица, g_n' – матрица Якоби функции g, k_i – стадии метода, a, p_i, β_{4j} – числовые коэффициенты **(4)**.

Для исследования схемы **(4)** разложим стадии k_i в ряды Тейлора. Получим **(5a)**

$$y_{n+1} = y_n + (p_1 + p_2 + p_3 + p_4)h\varphi_n + (p_2 + p_3)hg_n +$$
$$+(\beta_{41} + \beta_{42} + \beta_{43})p_4 h^2 \varphi'_n \varphi_n + (\beta_{42} + \beta_{43})p_4 h^2 \varphi'_n g_n +$$
$$+a(p_2 + 2p_3)h^2 g'_n \varphi_n + a(p_2 + 2p_3)h^2 g'_n g_n + O(h^3). \quad (5a)$$

Ряд Тейлора для точного решения $y(t_{n+1})$ имеет вид **(5b)**

$$y(t_{n+1}) = y(t_n) + h(\varphi + g) +$$
$$+0.5h^2(\varphi'\varphi + \varphi' g + g'\varphi + g'g) + O(h^3). \quad (5b)$$

Сравнивая полученные ряды, получим условия второго порядка точности схемы **(4)**, то есть **(5c)**

$$\mathbf{1)}\ p_1 + p_2 + p_3 + p_4 = 1,\ \mathbf{2)}\ p_2 + p_3 = 1,$$

$$\mathbf{3)}\ a(p_2 + 2p_3) = 0.5,\ \mathbf{4)}\ (\beta_{41} + \beta_{42} + \beta_{43})p_4 = 0.5,\quad \mathbf{(5c)}$$

$$\mathbf{5)}\ (\beta_{42} + \beta_{43})p_4 = 0.5.$$

Отсюда следует **(5d)**

$$\beta_{41} = 0,\ p_2 = a^{-1}(4a-1)/2,\ p_3 = a^{-1}(1-2a)/2,$$

$$p_1 + p_4 = 0,\ (\beta_{42} + \beta_{43})p_4 = 1/2.\quad \mathbf{(5d)}$$

Исследуем устойчивость **(4)**. Применение уравнения **(6)**

$$y' = \lambda y \text{ **с комплексным** } \lambda,\ \Re(\lambda) < 0,\quad \mathbf{(6)}$$

в данном случае неправомерно, поскольку **в этом случае теряется смысл в разделении правой части системы дифференциальных уравнений на жесткую и нежесткую часть.**

Поэтому в системе **(3)**

$$y' = [f(t,y) - By] + By \qquad \mathbf{(3)}$$

положим

$$\varphi(y) = \lambda_1 y,\ g(y) = \lambda_2 y, \qquad \mathbf{(7)}$$

где λ_1 и λ_2 есть произвольные комплексные числа, причем $\Re(\lambda_2) < 0$. Смысл λ_1 и λ_2 – некоторые собственные числа матриц Якоби функций $\varphi(y)$ и $g(y)$, соответственно.

Применяя схему **(4)** для решения задачи **(8)**

$$y' = \lambda_1 y + \lambda_2 y,\ y(0) = y_0,\ t \geq 0 \qquad \textbf{(8)}$$

и обозначая $x=h\lambda_1$ и $z=h\lambda_2$, имеем

$$y_{n+1}=Q(x,y)y_n, \qquad \textbf{(9)}$$

где функция устойчивости $Q(x,y)$ имеет вид **(9a)**

$$Q(x,z) = \{1+(1-2a)z + x +$$
$$+[-2ap_1 - ap_2 + (\beta_{42} + \beta_{43} - 2a)p_4]xz + 0.5x^2 -$$
$$-a\beta_{42}p_4x^2z + [a^2p_1 + a^2p_4 - a\beta_{42}p_4]xz^2 +$$
$$+(a^2 - ap_2)z^2\}/(1-az)^2. \qquad \textbf{(9a)}$$

Необходимым условием L-устойчивости численной формулы **(4)** относительно функции $g(y)=\lambda_2 y$ является выполнение соотношения $Q(x,z)\to 0$ при $z\to -\infty$. Из вида $Q(x,z)$ следует, что это требование будет выполнено, если **(9b)**

$$p_2 = a \text{ и } \beta_{42} = 0. \qquad \textbf{(9b)}$$

В результате получим набор коэффициентов схемы **(4)** второго порядка точности вида **(10)**

$$\beta_{41} = \beta_{42} = 0,\ p_2 = a,\ p_3 = 1-a,\ p_4 = -p_1 = \frac{1}{2\beta_{43}^{-1}}, \textbf{(10)}$$

где β_{43} – свободный параметр, а коэффициент a есть корень уравнения **(11a)**

$$a^2 - 2a + 0.5 = 0. \qquad \textbf{(11a)}$$

Теперь функция устойчивости $Q(x,z)$ имеет вид **(11b)**

$$Q(x,z)=\frac{1+x+0.5x^2+(1-2a)z+(1-2a)xz}{(1-az)^2} \qquad \textbf{(11b)}$$

Заметим, что если $\varphi(y)\equiv 0$, то схема **(4)** с коэффициентами **(10)** совпадает с L-устойчивым **(2,1)**-методом **(12a)**

$$y_{n+1}=y_n+ak_2+(1-a)k_3 \qquad \textbf{(12a)}$$

функция устойчивости $Q(0,z)$ которого имеет вид **(12b)**

$$Q(0,z)=\frac{1+(1-2a)z}{(1-az)^2}, \qquad \textbf{(12b)}$$

а локальная ошибка δ_n — **(12c)**

$$\delta_n=(a-\frac{1}{3})h^3g'^2g+\frac{1}{6}h^3g''g^2+O(h^4). \qquad \textbf{(12c)}$$

Уравнение устойчивости **(11a)**

$$a^2-2a+0.5=0, \qquad \textbf{(11a)}$$

$$a_1=1-0.5\sqrt{2} \text{ и } a_2=1+0.5\sqrt{2}$$

имеет два вещественных. Выберем $a=a_1$, потому что в этом случае меньше коэффициент в главном члене локальной ошибки **(2,1)**-схемы.

Если **$g(y)\equiv 0$**, , то численная формула **(4)** вырождается в явный двухстадийный метод типа Рунге-Кутта вида **(13a)**

$$y_{n+1} = y_n + (1 - 0.5/\beta_{43})k_1 + 0.5k_4/\beta_{43}. \qquad \textbf{(13a)}$$

Локальную ошибку δ_n схемы **(13a)** можно записать в виде **(13b)**

$$\delta_n = \frac{1}{6}h^3\varphi'^2\varphi + (\frac{1}{6} - \frac{1}{4}\beta_{43}^{-1})h^3\varphi''\varphi^2 + O(h^4). \qquad \textbf{(13b)}$$

Отсюда следует, что локальная ошибка явной формулы будет минимальной, если β_{43}**=2/3**.

Имеем коэффициенты схемы **(4)** второго порядка **(13c)**

$$a = 1 - \frac{\sqrt{2}}{2},\ \beta_{41} = \beta_{42} = 0,\ \beta_{43} = \frac{2}{3},$$

$$p_4 = -p_1 = \frac{3}{4},\ p_2 = a,\ p_3 = 1 - a. \qquad \textbf{(13c)}$$

КОНТРОЛЬ ТОЧНОСТИ ВЫЧИСЛЕНИЙ

Контроль вычислений схемы **(4)** будем осуществлять с помощью метода первого порядка точности вида **(14)**

$$y_{n+1,1} = y_n + h[\varphi(y_n) + g(y_n)]. \qquad (14)$$

Ошибку ε_n схемы **(4)** можно вычислить по формуле **(15)**

$$\varepsilon_n = y_{n+1} - y_{n+1,1}. \qquad (15)$$

Подчеркнем особенность этой оценки. В силу L-устойчивости схемы **(4)** следует, что для функции устойчивости $Q(x,z)$ выполняется $Q(x,z)\to 0$ при $z\to-\infty$. Так как для точного решения **(16)**

$$y(t_{n+1}) = \exp(x+z)\,y(t_n) \qquad (16)$$

задачи **(6)**

$$y' = \lambda y \text{ с комплексным } \lambda,\ \Re(\lambda) < 0, \qquad (6)$$

выполняется аналогичное свойство, то естественным будет требование стремления к нулю оценки ошибки ε_n при $z\to-\infty$. **Однако для построенной оценки имеем** $\varepsilon_n = O(z)$.

Поэтому с целью исправления асимптотического поведения, вместо ε_n рассмотрим оценки $\varepsilon_n(j_n)$ вида **(17)**

$$\varepsilon_n(j_n) = D^{1-j_n}\varepsilon_n,\ 1 \le j_n \le 3. \qquad (17)$$

В смысле главного члена, то есть первого члена при разложении ошибок в ряды Тейлора по степеням h, оценки ε_n и $\varepsilon_n(j_n)$ совпадают при любом значении j_n, причем $\varepsilon_n(j_n)\to 0$ при $z\to-\infty$.

Теперь для контроля точности вычислений можно применять неравенство **(18)**

$$\| \boldsymbol{\varepsilon}_n(\boldsymbol{j}_n) \| \le \boldsymbol{\varepsilon}, \; 1 \le \boldsymbol{j}_n \le 3, \qquad \textbf{(18)}$$

где ε – точность расчетов, $\|\bullet\|$ – некоторая норма в $\boldsymbol{R}^N$.

Применение $\varepsilon_n(j_n)$ вместо ε_n не приводит к существенному увеличению вычислительных затрат. При $z \to 0$ оценка $\varepsilon_n(1)=\varepsilon_n$ правильно отражает поведение ошибки и нет смысла проверять при других значениях j_n.

При резком увеличении шага поведение ε_n может оказаться неудовлетворительным, что проявляется в неоправданном уменьшении шага и повторных вычислениях решения. Поэтому при реализации алгоритма интегрирования неравенство для контроля точности используется следующим образом.

При каждом фиксированном n выбирается наименьшее значение j_n, при котором выполняется неравенство. Если оно не выполняется ни при каком j_n , то шаг уменьшается и решение вычисляется повторно.

ЧИСЛЕННАЯ СХЕМА ДЛЯ НЕАВТОНОМНЫХ ЗАДАЧ

Рассмотрим задачу Коши для неавтономной системы **(19)**

$$y' = \varphi(t, y) + g(t, y),\ y(t_0) = y_0,\ t_0 \le t \le t_k. \quad (19)$$

Снова предполагаем, что вся жесткость сосредоточена в функции $g(t,y)$, а $\varphi(t,y)$ есть нежесткая часть. Для численного решения **(19)** рассмотрим метод вида **(20)**

$$y_{n+1} = y_n + \sum_{i=1}^{4} p_i k_i,\ D = E - ahg'_n,\ k_1 = h\varphi(t_n, y_n),$$

$$Dk_2 = h[\varphi(t_n, y_n) + g(t_n + ch, y_n)],\ Dk_3 = k_2, \quad (20)$$

$$k_4 = h\varphi(t_n + (\beta_{41} + \beta_{42} + \beta_{43})h, y_n + \beta_{41}k_1 + \beta_{42}k_2 + \beta_{43}k_3)$$

Разложим стадии метода **(20)** в ряды Тейлора и подставим в первую формулу **(20)**, получим **(21)**

$$y_{n+1} = y_n + (p_1 + p_2 + p_3 + p_4)h\varphi_n + (p_2 + p_3)hg_n +$$

$$+(\beta_{41} + \beta_{42} + \beta_{43})p_4 h^2 \varphi'_{tn} + c(p_2 + p_3)g'_{tn} +$$

$$+(\beta_{41} + \beta_{42} + \beta_{43})p_4 h^2 \varphi'_n \varphi_n + (\beta_{42} + \beta_{43})p_4 h^2 \varphi'_n g_n +$$

$$+a(p_2 + 2p_3)h^2 g'_n \varphi_n + a(p_2 + 2p_3)h^2 g'_n g_n + O(h^3). \quad (21)$$

Представление точного решения $y(t_{n+1})$ имеет вид **(22)**

$$y(t_{n+1}) = y(t_n) + h(\varphi + g) +$$

$$+\frac{1}{2}h^2(\varphi'_t + g'_t + \varphi'\varphi + \varphi'_y g + g'_y \varphi + g'g) + O(h^3). \quad (22)$$

Сравнивая полученные ряды, получим условия второго порядка точности **(23)**

$$p_1 + p_2 + p_3 + p_4 = 1,\ p_2 + p_3 = 1,$$
$$(\beta_{41} + \beta_{42} + \beta_{43})p_4 = 0.5,\ c(p_2 + p_3) = 0.5, \qquad (23)$$
$$(\beta_{42} + \beta_{43})p_4 = 0.5,\ a(p_2 + 2p_3) = 0.5.$$

Отсюда сразу следует c**=0.5.** Теперь рассуждая по аналогии с исследованием схемы **(4)**, получим коэффициенты численной формулы **(20)**, которые имеют вид **(24)**

$$a = 1 - \frac{\sqrt{2}}{2},\ \beta_{41} = \beta_{42} = 0,\ c = \frac{1}{2},\ \beta_{43} = \frac{2}{3},$$
$$p_1 = -\frac{3}{4},\ p_2 = a,\ p_3 = 1 - a,\ p_4 = \frac{3}{4}. \quad (24)$$

Неравенство для контроля точности вычислений построим по аналогии со схемой **(4)**, где в оценке ε_n используется приближенное решение по методу первого порядка **(25)**

$$y_{n+1,1} = y_n + h[\varphi(t_n, y_n) + g(t_n + 0.5h, y_n)]. \qquad (25)$$

АНАЛИЗ РЕЗУЛЬТАТОВ РАСЧЕТОВ

Все рассматриваемые примеры приводились к виду **(2)**

$$y' = [\,f(t,y) - By] + By,\ \ y(t_0) = y_0,\ \ t_0 \le t \le t_k, \quad (2)$$

Вычисления осуществлялись с требуемой точностью $\varepsilon=10^{-2}$. Расчеты проводились с двойной точностью.

Схема **(4)** имеет второй порядок точности и поэтому проводить с ее помощью расчеты с более высокой точностью нецелесообразно.

В расчетах левая часть неравенства для контроля точности вычислялась по формуле **(26)**

$$\|\,\varepsilon_n(j_n)\,\| = \max_{1\le i\le N}\left\{\frac{|\,\varepsilon_n^i(j_n)\,|}{|\,y_n^i\,| + r}\right\}. \quad (26)$$

где i – номер компоненты, r – положительный параметр. Если по i-й компоненте решения выполняется неравенство $|y_n^i|<r$, то контролируется абсолютная ошибка εr, в противном случае – относительная ошибка ε. В расчетах параметр r выбирался таким образом, чтобы по всем компонентам решения фактическая точность была не хуже задаваемой.

Ниже через **is, if, idec, isol** обозначены, соответственно, суммарное число шагов интегрирования, правых частей системы **(1)**, декомпозиций матрицы Якоби и число обратных ходов в методе Гаусса.

Пример 1.

$$y_2' = -0.013y_1 - 1000y_1y_3,\; y_2' = -2500y_2y_3,$$
$$y_3' = -0.013y_1 - 1000y_1y_3 - 2500y_2y_3, \qquad (27)$$
$$t \in [0,50],\; y_1(0) = y_2(0) = 1,\; y_3(0) = 0,\; h_0 = 2.9\cdot 10^{-4}.$$

Решение задачи **(27)** осуществлялось методом **(4)** с диагональной аппроксимацией матрицы Якоби.

Так как в этом случае вычислительные затраты метода (4) практически такие же, как и в явных методах, то сравнение эффективности проводилось с известным явным методом Мерсона четвертого порядка точности.

Для вычисления приближенного решения построенным алгоритмом ASODE2 **потребовалось 687 шагов, остальные затраты вычисляются из вида схемы (4).** Для решения задачи **(27)** **методу Мерсона потребовалось 400 627** вычислений правой части.

Пример 2.

$$\boldsymbol{y}_1' = -55\boldsymbol{y}_1 + 65\boldsymbol{y}_2 - \boldsymbol{y}_1\boldsymbol{y}_2,$$
$$\boldsymbol{y}_2' = 0.0785(\boldsymbol{y}_1 - \boldsymbol{y}_2),\ \boldsymbol{y}_3' = 0.1\boldsymbol{y}_1, \qquad (28)$$
$$\boldsymbol{t} \in [0,500],\ \boldsymbol{y}_1(0) = \boldsymbol{y}_2(0) = 1,\ \boldsymbol{y}_3(0) = 0,\ \boldsymbol{h}_0 = 2\cdot 10^{-2}.$$

Решение задачи **(28)** осуществлялось методом **(4)** с диагональной аппроксимацией матрицы Якоби.

Приближенное решение **алгоритмом ASODE2 вычислено за 4 953 шага.** Для решения данной задачи **методу Мерсона потребовалось 80 713 вычислений правой части.**

Пример 3.

$$\boldsymbol{y}_1' = 77.27(\boldsymbol{y}_1(1 - 8.375\cdot 10^{-6}\boldsymbol{y}_1 - \boldsymbol{y}_2) + \boldsymbol{y}_2),$$
$$\boldsymbol{y}_2' = (\boldsymbol{y}_3 - (1 + \boldsymbol{y}_1)\boldsymbol{y}_2)/77.27,\ \boldsymbol{y}_3' = 0.161(\boldsymbol{y}_1 - \boldsymbol{y}_3), \quad (29)$$
$$\boldsymbol{t} \in [0,360],\ \boldsymbol{y}_1(0) = 1,\ \boldsymbol{y}_2(0) = 2,\ \boldsymbol{y}_3(0) = 3,\ \boldsymbol{h}_0 = 10^{-6}.$$

Решение задачи **(29)** осуществлялось методом **(4)** с диагональной аппроксимацией матрицы Якоби. Приближенное решение **алгоритмом ASODE2 вычислено за 19 964 шага.** Для решения данной задачи **методу Мерсона потребовалось 2 3700 664 вычислений правой части.**

ЗАКЛЮЧЕНИЕ

Предложенный алгоритм интегрирования создавался для численного решения задач механики сплошной среды после дискретизации по пространству методом конечных элементов или с помощью конечных разностей. В этом случае в **(3)**

$$y' = \varphi(y) + g(y), \; y(t_0) = y_0, \; t_0 \le t \le t_k, \quad \textbf{(3)}$$

разделение на функции $g(y)$ и $\varphi(y)$ происходит естественно — $g(y)$ **есть симметричная часть,** описываемая оператором дифференцирования второго порядка, а $\varphi(y)$ **есть несимметричная часть** (**конвективные слагаемые**), описываемые оператором первого порядка.

При реализации формулы **(4)** необходимо дважды решать линейную систему алгебраических уравнений. В задачах механики сплошной среды эффективность алгоритма интегрирования может быть достигнута за счет специальных методов решения линейных систем с симметричной матрицей, которая во многих случаях положительно определенная.

Схему **(4)** можно применять также для решения локально-неустойчивых задач, причем $\varphi(y)$ в этом случае отвечает за собственные числа матрицы Якоби с положительными вещественными частями. В отличие от L-устойчивых методов, у которых область неустойчивости обычно небольшая и которые являются L-устойчивыми не только в левой, но и в правой полуплоскости плоскости $\{h\lambda\}$, явные методы типа Рунге-Кутта являются неустойчивыми практически во всей правой полуплоскости и поэтому более предпочтительны при определении неустойчивого решения.

Для локально-неустойчивых задач в ряде случаев разделение правой части системы обыкновенных дифференциальных уравнений на функции $\varphi(y)$ и $g(y)$ из физических соображений тоже не вызывает особых трудностей.

Приведенные результаты численных экспериментов не ориентированы на решение задач механики сплошной среды или локально-неустойчивых задач, а направлены на исследование возможностей алгоритма интегрирования при решении некоторых общепринятых тестовых примеров. Тестовые примеры выбирались таким образом, чтобы продемонстрировать разные нюансы работы алгоритма интегрирования.

Цель расчетов заключалась в проверке работоспособности алгоритма с переменным шагом и с замораживанием матрицы Якоби, надежности неравенства для контроля точности вычислений, а также в исследовании возможности расчетов с диагональной аппроксимацией матрицы Якоби.

В последнем случае вычислительные затраты на шаг в явных методах и построенном алгоритме различаются незначительно. В частности из анализа результатов расчетов жестких задач следует, что в случае невозможности применения методов с неограниченной областью устойчивости, **алгоритм (4) существенно эффективнее метода Мерсона — наиболее распространенного среди явных численных схем типа Рунге-Кутта.**

СПАСИБО

ЗА

ВНИМАНИЕ