

ЗАМОРАЖИВАНИЕ МАТРИЦЫ ЯКОБИ В МЕТОДАХ ТИПА РОЗЕНБРОКА*

Е. А. НОВИКОВ, А. Л. ДВИНСКИЙ

Красноярский государственный технический университет, Россия

e-mail: novikov@icm.krasn.ru, dvinski@inetik.ru

The problem of freezing of the Jacobi matrix in Rosenbrock-type methods is investigated. Numerical tests confirm that the Jacobi matrix freezing improve the efficiency of calculations.

Введение

Во многих приложениях возникает необходимость численного решения задачи Коши для жестких систем обыкновенных дифференциальных уравнений. Стремление ко все более точному описанию физических процессов приводит к постоянному росту размерности и жесткости соответствующих задач. В результате требования, предъявляемые к алгоритмам интегрирования, постоянно повышаются.

При решении жестких задач широкое распространение получили методы типа Розенброка [1]. Это связано с их хорошими свойствами точности и устойчивости, а также с простотой реализации на ЭВМ. В данных методах матрица Якоби внесена непосредственно в численную формулу, и поэтому вопрос ее использования на нескольких шагах интегрирования нуждается в специальном изучении.

При решении жестких задач на эффективность численной схемы существенно влияют свойства устойчивости не только основной, но и промежуточных или внутренних численных формул [2–5]. За счет согласования свойств устойчивости основной и промежуточных численных схем можно повысить эффективность алгоритма интегрирования.

Здесь доказана теорема о максимальном порядке точности методов типа Розенброка при некоторых аппроксимациях матрицы Якоби. Построена L -устойчивая схема максимального порядка точности с L -устойчивой внутренней схемой, которую можно использовать с замораживанием как аналитической, так и численной матрицы Якоби. Получено неравенство для контроля точности вычислений, не приводящее к росту вычислительных затрат. Приведены результаты численных расчетов, подтверждающие повышение эффективности метода за счет замораживания матрицы Якоби.

*Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (грант № 05-01-00579-а).

© Институт вычислительных технологий Сибирского отделения Российской академии наук, 2005.

1. Методы типа Розенброка

Рассмотрим задачу Коши для системы обыкновенных дифференциальных уравнений

$$y' = f(y), \quad y(t_0) = y_0, \quad t_0 \leq t \leq t_k, \quad (1)$$

где y и f — N -мерные вещественные вектор-функции; t — независимая переменная. Для решения (1) применим методы типа Розенброка [1], которые имеют вид

$$y_{n+1} = y_n + \sum_{i=1}^m p_i k_i, \quad D_n = E - ahf'_n, \\ D_n k_i = hf \left(y_n + \sum_{j=1}^{i-1} \beta_{ij} k_j \right), \quad (2)$$

где E — единичная матрица; $f'_n = \partial f(y_n)/\partial y$ — матрица Якоби системы (1); h — шаг интегрирования; k_i , $1 \leq i \leq m$, — стадии метода; a, p_i, β_{ij} — вещественные константы, определяющие свойства точности и устойчивости численных формул (2).

Пусть при реализации (2) применяется замораживание матрицы Якоби. Это означает, что матрица D_n вычисляется по формуле вида $D_n = E - ahf'_{n-i}$, где i есть число шагов с замороженной матрицей. Нетрудно видеть, что в этом случае имеет место соотношение $f'_{n-i} = f'_n - ih(f''f)_\zeta$, где матрица $(f''f)_\zeta$ вычисляется в некоторой точке ζ . Аналогичное соотношение можно получить при численном вычислении матрицы Якоби [6]. Поэтому далее будем предполагать, что при реализации (2) используется матрица

$$D_n = E - ahA_n, \quad (3)$$

где A_n представима в виде

$$A_n = f'_n + hB_n, \quad (4)$$

а матрица B_n не зависит от размера шага интегрирования. Данное предположение позволяет применять методы (2) с замораживанием как аналитической, так и численной матрицы Якоби, а также допускает некоторые другие виды ее аппроксимации.

Теперь исследуем влияние аппроксимации матрицы Якоби на порядок точности методов типа Розенброка. Предварительно докажем следующую лемму.

Лемма. Пусть при реализации методов (2) применяется матрица (3). Тогда стадии k_i , $i \leq i \leq m$, представимы в виде

$$k_i = hf_n + (a + \alpha_i)h^2 f'_n f_n + (a^2 + 2a\alpha_i + \alpha'_i)h^3 f'^2_n f_n + \\ + 0.5\alpha_i^2 h^3 f''_n f_n^2 + ah^3 B_n f_n + O(h^4), \quad (5)$$

где

$$\alpha_i = \sum_{j=1}^{i-1} \beta_{ij}, \quad 2 \leq i \leq m, \quad \alpha_1 = 0, \\ \alpha'_i = \sum_{j=1}^{i-1} \beta_{ij} \alpha_j, \quad 3 \leq i \leq m, \quad \alpha'_1 = \alpha'_2 = 0,$$

а элементарные дифференциалы

$$f_n = f(y_n), \quad f'_n f_n = \frac{\partial f(y_n)}{\partial y} f(y_n), \quad f''_n f_n^2 = \left(\frac{\partial^2 f(y_n)}{\partial y^2} f(y_n), f(y_n) \right)$$

вычислены на приближенном решении y_n .

Доказательство проведем методом математической индукции. Нетрудно видеть, что матрица D_n^{-1} представима в виде

$$D_n^{-1} = E + ahf'_n + a^2h^2f''_n + ah^2B_n + O(h^3).$$

Тогда при $i = 1$ соотношение (5) очевидно. Пусть k_i имеет вид (5). Тогда при $j = i + 1$ имеем

$$D_n k_j = hf \left(y_n + \sum_{l=1}^{j-1} \beta_{jl} k_l \right). \quad (6)$$

Разложим правую часть (6) в ряд Тейлора по степеням h до членов с h^3 включительно. Получим

$$\begin{aligned} hf \left(y_n + \sum_{l=1}^{j-1} \beta_{jl} k_l \right) &= hf \left[y_n + \left(\sum_{l=1}^{j-1} \beta_{jl} \right) hf_n + \left(\sum_{l=1}^{j-1} \beta_{jl} (a + \alpha_l) \right) h^2 f'_n f_n + O(h^3) \right] = \\ &= hf_n + \left(\sum_{l=1}^{j-1} \beta_{jl} \right) h^2 f'_n f_n + \left(a \sum_{l=1}^{j-1} \beta_{jl} + \sum_{l=1}^{j-1} \beta_{jl} \alpha_l \right) h^3 f''_n f_n + 0.5 \left(\sum_{l=1}^{j-1} \beta_{jl} \right)^2 h^3 f''_n f_n^2 + \\ &+ O(h^4) = hf_n + \alpha_j h^2 f'_n f_n + (a\alpha_j + \alpha'_j) h^3 f''_n f_n + 0.5\alpha_j^2 h^3 f''_n f_n^2 + O(h^4). \end{aligned}$$

Применяя слева к полученному соотношению D_n^{-1} , получим

$$\begin{aligned} k_j &= hf_n + (a + \alpha_j) h^2 f'_n f_n + (a^2 + 2a\alpha_j + \alpha'_j) h^3 f''_n f_n + \\ &+ 0.5\alpha_j^2 h^3 f''_n f_n^2 + ah^3 B_n f_n + O(h^4). \end{aligned}$$

Отсюда, учитывая, что $j = i + 1$, следует доказательство леммы. \square

Теорема. Пусть при реализации методов Розенброка применяется матрица (3). Тогда максимальный порядок точности численных схем (2) равен двум.

Доказательство. Для доказательства подставим (5) в первую формулу (2). Получим

$$\begin{aligned} y_{n+1} &= y_n + \left(\sum_{i=1}^m p_i \right) hf_n + \left(\sum_{i=1}^m p_i (a + \alpha_i) \right) h^2 f'_n f_n + \\ &+ \left(\sum_{i=1}^m p_i (a^2 + 2a\alpha_i + \alpha'_i) \right) h^3 f''_n f_n + 0.5 \left(\sum_{i=1}^m p_i \alpha_i^2 \right) h^3 f''_n f_n^2 + \\ &+ a \left(\sum_{i=1}^m p_i \right) h^3 B_n f_n + O(h^4). \end{aligned} \quad (7)$$

Здесь элементарные дифференциалы вычислены на приближенном решении y_n . Представление точного решения $y(t_{n+1})$ в окрестности точки t_n в виде ряда Тейлора по степеням h до членов с h^3 включительно имеет вид

$$y(t_{n+1}) = y(t_n) + hf_n + \frac{h^2}{2} f'_n f_n + \frac{h^3}{6} (f''_n f_n + f''_n f_n^2) + O(h^4), \quad (8)$$

где элементарные дифференциалы вычислены на точном решении $y(t_n)$.

Сравнивая (7) и (8) до членов с h^3 включительно при условии $y_n = y(t_n)$, получим условия третьего порядка точности схемы (2), т. е.

$$\begin{aligned} \sum_{i=1}^m p_i &= 1, & \sum_{i=1}^{m-1} \alpha_i p_i &= 0.5 - a, \\ 2a \sum_{i=1}^{m-1} \alpha_i p_i + \sum_{i=1}^{m-2} \alpha'_i p_i &= 1/6 - a^2, \\ \sum_{i=1}^{m-1} \alpha_i^2 p_i &= 1/3, & a \sum_{i=1}^m p_i &= 0. \end{aligned} \quad (9)$$

Отсюда следует, что последнее уравнение (9) противоречит требованию первого порядка точности, что завершает доказательство. \square

2. Схема второго порядка точности

Рассмотрим двухстадийную схему типа Розенброка при условиях (3) и (4), т. е.

$$y_{n+1} = y_n + p_1 k_1 + p_2 k_2, \quad D_n k_1 = hf(y_n), \quad D_n k_2 = hf(y_n + \beta k_1). \quad (10)$$

Из (9) следует, что (10) будет иметь второй порядок точности, если

$$p_1 + p_2 = 1, \quad \beta p_2 = 0.5 - a. \quad (11)$$

Исследуем устойчивость (10) на линейном скалярном уравнении

$$y' = \lambda y, \quad \text{Re}(\lambda) < 0. \quad (12)$$

Применяя (10) для решения (12) и учитывая (11), получим $y_{n+1} = Q(x)y_n$, где

$$Q(x) = \frac{1 + (1 - 2a)x + (a^2 - 2a + 0.5)x^2}{(1 - ax)^2}, \quad x = \lambda h. \quad (13)$$

Отсюда следует, что для L -устойчивости (10) необходимо выполнение равенства

$$a^2 - 2a + 0.5 = 0. \quad (14)$$

Далее, из анализа результатов расчетов следует, что недостаточно хорошие свойства устойчивости внутренней или промежуточной схем

$$y_{n+\beta} = y_n + \beta k_1 \quad (15)$$

приводят к понижению эффективности алгоритма интегрирования. Так как при условии (14) основная схема L -устойчивая, потребуем аналогичного свойства от (15). Применяя ее для решения (12), получим $y_{n+\beta} = Q_\beta(x)y_n$, где $Q_\beta = [1 + (\beta - a)x]/(1 - ax)$. Отсюда следует, что внутренняя схема (15) будет L -устойчивой, если $\beta = a$. Учитывая полученное

соотношение и равенства (11) и (14), локальную ошибку δ_n схемы (10) можно записать в виде

$$\delta_n = \left(a - \frac{1}{3}\right) h^3 f_n'^2 f_n + \frac{9a-1}{12} h^3 f_n'' f_n^2 + ah^3 B_n f_n + O(h^4).$$

Уравнение (14) имеет два корня $a_1 = 1 - \sqrt{2}/2$ и $a_2 = 1 + \sqrt{2}/2$. Выберем $a = 1 - \sqrt{2}/2$, потому что в этом случае меньше коэффициент в главном члене локальной ошибки. В результате получим набор коэффициентов

$$p_1 = \beta = a = 1 - \frac{\sqrt{2}}{2}, \quad p_2 = 1 - a, \tag{16}$$

при которых численная формула (10) имеет второй порядок точности, является L -устойчивой и обладает L -устойчивой промежуточной схемой.

3. Алгоритм интегрирования

Контроль точности (10), (16) будем осуществлять методом первого порядка точности

$$y_{n+1,1} = y_n + k_1.$$

Тогда в неравенстве для контроля точности вычислений можно использовать оценку ошибки ε_n вида

$$\varepsilon_n = y_{n+1} - y_{n+1,1} = (1 - a)(k_2 - k_1). \tag{17}$$

Подчеркнем важную особенность построенной оценки ошибки. В силу L -устойчивости из (13) и (14) следует, что для функции устойчивости $Q(x)$ выполняется $Q(x) \rightarrow 0$ при $x \rightarrow -\infty$. Так как для точного решения $y(t_{n+1}) = \exp(x)y(t_n)$ задачи (12) выполняется аналогичное свойство, естественным будет требование стремления к нулю оценки ошибки (17) при $x \rightarrow -\infty$. Однако для построенной оценки имеем $\varepsilon_n = O(1)$. Поэтому с целью исправления асимптотического поведения вместо ε_n рассмотрим оценки $\varepsilon_n(j_n)$ вида

$$\varepsilon_n(j_n) = D^{1-j_n} \varepsilon_n, \quad 1 \leq j_n \leq 2. \tag{18}$$

Нетрудно видеть, что в смысле главного члена, т. е. 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 $\varepsilon_n(2) \rightarrow 0$ при $x \rightarrow -\infty$. Теперь для контроля точности вычислений можно применять неравенство

$$\|\varepsilon_n(j_n)\| \leq \varepsilon_n, \quad 1 \leq j_n \leq 2. \tag{19}$$

Отметим, что применение $\varepsilon_n(j_n)$ вместо ε_n не приводит к существенному увеличению вычислительных затрат. При $x \rightarrow 0$ оценка $\varepsilon_n(1) = \varepsilon_n$ правильно отражает поведение ошибки и нет смысла проверять (19) 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 (19) используется следующим образом. При каждом фиксированном n выбирается наименьшее значение j_n , при котором выполняется неравенство (19). Если оно не выполняется ни при каком j_n , то шаг уменьшается и решение вычисляется повторно.

Замораживание матрицы Якоби, т. е. применение матрицы $D_n = E - ahA_n$ на нескольких шагах интегрирования, проводилось по следующему правилу. Если матрица Якоби не пересчитывалась, то для сохранения устойчивости численной схемы величина шага интегрирования тоже не менялась. Попытка применения старой матрицы осуществлялась после каждого успешного шага интегрирования. Три причины приводили к размораживанию: 1) нарушение неравенства для контроля точности вычислений; 2) превышение количества шагов с замороженной матрицей числа q_f ; 3) превышение прогнозируемого шага интегрирования последнего успешного шага в q_h раз. Параметры q_f и q_h можно использовать для настройки метода на решение конкретных задач. Если $q_f \rightarrow \infty$ и $q_h \rightarrow \infty$, то число шагов с одной матрицей Якоби возрастает. Если $q_f = 0$ и $q_h = 0$, то замораживание матрицы не происходит. Поэтому в случае большой размерности системы обыкновенных дифференциальных уравнений имеет смысл выбирать q_f и q_h достаточно большими.

4. Результаты расчетов

Расчеты проводились с требуемой точностью $\varepsilon = 10^{-2}$ на IBM PC Athlon(tm)XP 2000+ с двойной точностью. Схема (10) имеет второй порядок точности, и поэтому проводить с ее помощью расчеты с более высокой точностью нецелесообразно. В конкретных расчетах левая часть неравенства (19) вычислялась по формуле

$$\|\varepsilon_n(j_n)\| = \max_{1 \leq i \leq N} \frac{|\varepsilon_n^i(j_n)|}{|y_n^i| + r}, \quad (20)$$

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. Под точным решением понималось решение, полученное методами [7] с требуемой точностью $\varepsilon = 10^{-6}$.

Эффективность построенного алгоритма без замораживания (ROZ-2) и с замораживанием (ROZ-2-FREEZ) матрицы Якоби сравнивалась с эффективностью метода Гира (LSODE) [8] на десяти тестовых примерах [9] из области химической кинетики. В ROZ-2-FREEZ применялись константы $q_f = 10$ и $q_h = 2$.

На вычисление всех примеров алгоритму ROZ-2 потребовалось $if = 832$ вычислений правой части и $jac = 323$ вычислений матрицы Якоби задачи (1). Для алгоритма ROZ-2-FREEZ затраты $if = 824$ и $jac = 159$, для метода Гира $is = 729$, $jac = 170$. Из анализа результатов расчетов следует, что за счет замораживания матрицы Якоби число ее вычислений сократилось вдвое, в то время как число вычислений правой части у алгоритмов ROZ-2 и ROZ-2-FREEZ фактически совпадает. Вычислительные затраты методов LSODE и ROZ-2-FREEZ практически одинаковые.

Список литературы

- [1] ROSENBROCK H.H. Some general implicit processes for the numerical solution of differential equations // Computer. 1963. N 5. P. 329–330.
- [2] ХАЙРЕР Э., НЕРСЕТТ С., ВАННЕР Г. Решение обыкновенных дифференциальных уравнений. Нежесткие задачи: Пер. с англ. М.: Мир, 1990. 512 с.

- [3] ХАЙРЕР Э., ВАННЕР Г. Решение обыкновенных дифференциальных уравнений. Жесткие и дифференциально-алгебраические задачи: Пер. с англ. М.: Мир, 1999. 685 с.
- [4] НОВИКОВ Е.А. Явные методы для жестких систем. Новосибирск: Наука, 1997. 195 с.
- [5] ДВИНСКИЙ А.Л. Исследование (m, k) -методов с L -устойчивыми промежуточными схемами для решения жестких систем: Дис. ... канд. физ.-мат. наук. Красноярск, 2004. 150 с.
- [6] НОВИКОВ Е.А., ШИТОВ Ю.А. Алгоритм интегрирования жестких систем на основе (m, k) -метода второго порядка точности с численным вычислением матрицы Якоби. Красноярск, 1988 (Препр. ВЦ СО АН СССР. № 20).
- [7] НОВИКОВ Е.А., ШИТОВ Ю.А., ШОКИН Ю.И. Одношаговые безытерационные методы решения жестких систем // Докл. АН СССР. 1988. Т. 301, № 6. С. 1310–1314.
- [8] HINDMARSH A.C. ODEPACK, a systematized collection of ODE solvers. Lawrence Livermore National Laboratory, 1982 (Preprint UCRL-88007).
- [9] ENRIGHT W.H., HULL T.E. Comparing numerical methods for the solutions of systems of ODE's // BIT. 1975. N 15. P. 10–48.

Поступила в редакцию 2 ноября 2005 г.