УДК 519.62

В. В. БОБКОВ, Н. А. БОБКОВА

К ВОПРОСУ О ПОСТРОЕНИИ МЕТОДОВ ЧИСЛЕННОГО РЕШЕНИЯ ЖЕСТКИХ СИСТЕМ

Основное внимание нами будет уделено задаче с начальными условиями для системы обыкновенных дифференциальных уравнений

$$u' = f(t,u), \tag{1}$$

где $u = (u_1, \dots, u_n)^T$, $f = (f_1, \dots, f_n)^T$, $u_i = u_i(t)$, $f_i = f_i(t, u)$, $i = 1, 2, \dots, n$. Однако многое из изложенного ниже может быть использовано при конструировании разностных схем и в случае других эволюционных задач с дифференциальными операторами более сложной структуры, а также при разработке итерационных процессов для стационарных задач, привязанных к соответствующей системе вида (1) через идею установления решения во времени.

При построении вычислительных алгоритмов для решения исходной дифференциальной задачи обычно обязательным этапом является переход к аппроксимирующей ее разностной задаче, многие свойства которой (напр., устойчивость) часто рассматривают [1] как внутренние свойства разностной схемы. Очевидно, что замена дифференциальной модели разностной, не претерпевающей конструктивных изменений во времени, значительно снижает возможности так организованного численного моделирования. И если адекватность дифференциальной модели исследуемому процессу может сохранятся при этом на всем рассматриваемом промежутке, то для разностной модели (особенно в случае явных схем) такой адекватности часто удается достичь лишь при значениях шагов сетки, значительно меньших, чем те, которые являются естественными для наблюдения данной стадии процесса. Это особенно актуально в том случае, когда мы сталкиваемся с явлением жесткости [2] и вынуждены численно анализировать медленно изменяющуюся стадию процесса при наличии быстро затухающих возмущений. Существенная разномасштабность составляющих процесса обычно заставляет при выборе шага сетки ориентироваться на более «быстрые» составляющие и в той стадии процесса, когда их вкладом в пределах точности наблюдения уже можно было бы пренебречь. Более экономичным мог бы быть такой вычислительный алгоритм, в основе которого лежит требование хорошей аппроксимации только тех составляющих решения, которые наблюдаемы в пределах заданной точности, а для других составляющих обеспечен лишь некоторый уровень их качественного поведения. При построении такого типа разностных схем, привязанных к интересующему нас решению, следует обеспечить выполнение ряда новых условий, на часть из которых было, например, обращено внимание в [3]. В качестве системы дифференциальных уравнений, на которую естественно ориентироваться при постановке такого рода дополнительных требований к конструируемым методам (в том числе и универсального назначения), может быть взята система

$$u' = Au + b \tag{2}$$

с постоянными матрицей A и вектором b. Такой выбор обусловлен не только тем, что проблема численного решения систем вида (2) представляет значительный самостоятельный интерес, а основные задачи линейной алгебры, построение итерационных процессов для которых может быть непосредственно связано с (2), наиболее часто встречаются в вычислительной практике, но и тем, что в общем случае системы (1) часто бывает оправданной предварительная аппроксимация на каждом шаге сетки исходной задачи последовательностью линейных задач (в сочетании с известным приемом «замораживания» коэффициентов, а также с оценкой области адекватности решений таких задач). Конструктивное обеспечение таких дополнительных требований согласованности дифференциальной и соответствующей разностной задач удобно проводить, скажем, с использованием идеи построения результирующего разностного оператора на вспомогательной сетке узлов [4] либо посредством корректировки спектральных свойств разностного оператора [5]. В последнем случае удается избежать трудоемкой операции умножения матриц, что особенно существенно в случае систем большой размерности. Данная работа выполнена в развитие [5].

Опираясь на прием пошагового выделения и точного обращения главной части исходного дифференциального оператора [6], применительно к (2) запишем интегральное соотношение:

$$u(t+\tau) = u(t) + \tau\rho(\mu\tau)[\mu u(t+\tau) + a]\exp(-\mu\tau) +$$
$$+ \int_t^{t+\tau} [Au(x) + b - \mu u(x) - a]\exp[\mu(t-x)]dx, \tag{3}$$

где $\rho(z) = (\exp z - 1)/z$, а правила выбора значений скалярного μ и векторного a параметров будут предопределяться соображениями, излагаемыми ниже. Точное соотношение (3), в отличие от используемого в [5], ориентировано на неявные методы.

Заменив в (3) интеграл простейшей квадратурной суммой с экспоненциальным весом и вводя корректирующий множитель σ (типа $\sigma = 1/(1+\alpha\tau)$ из [5] с числовым параметром $\alpha \geq 0$), приходим к приближенному равенству:

$$\hat{u} \approx u + \tau\rho(\mu\tau)\exp(-\mu\tau)[\mu\hat{u} + a + (Au + b - \mu u - a)\sigma], \tag{4}$$

где для компактности записи использованы обозначения $u = u(t)$, $\hat{u} = u(t+\tau)$. С (4) сопряжем метод вида:

$$\hat{y} = y + \tau\rho(\mu\tau)\exp(-\mu\tau)[\mu\hat{y} + a + (Ay + b - \mu y - a)\sigma], \tag{5}$$

зависящий от параметров μ, a и корректора σ. Здесь $y \approx u$, $\hat{y} \approx \hat{u}$, $\tau > 0$.

Учитывая скалярную неявность метода (5), его легко записать в форме

$$\hat{y} = Sy + g = y + Qy + g, \tag{6}$$

где

$$S = I + Q = \exp(\mu\tau)I + \tau\rho(\mu\tau)(A - \mu I)\sigma, \tag{7}$$

$$g = \tau\rho(\mu\tau)[a + (b - a)\sigma]. \tag{8}$$

Подчиним выбор σ требованию спектральной согласованности матриц S и $\exp(A\tau)$, при этом будем исходить, например, из следующих соображений.

Если в рассматриваемом n-мерном векторном пространстве решений системы (2) из собственных векторов матрицы A можно выбрать, например, ортонормированный базис $\xi^1, \xi^2, \ldots, \xi^n$, отвечающий собственным значениям $\lambda_1, \lambda_2, \ldots, \lambda_n$ соответственно, то любая траектория системы (2) с неособенной матрицей A может быть записана в виде:

$$u(t) = -A^{-1}b + \sum_{i=1}^{n} c_i\xi^i\exp(\lambda_i t) \tag{9}$$

со своими значениями коэффициентов $c_1, c_2, \ldots c_n$, определяемыми ее начальным состоянием. Применив метод (6), (7), (8) к системе (2) при $y = u(t)$ из (9), придем к равенству:

$$\hat{y} = g - SA^{-1}b + \sum_{i=1}^{n} c_i\xi^i\exp(\lambda_i t)s_i, \tag{10}$$

где

$$s_i = \exp(\mu\tau) + \tau\rho(\mu\tau)(\lambda_i - \mu)\sigma. \tag{11}$$

Обычно для явного типа методов применительно к (2) обременительные ограничения на шаг τ, связанные с обеспечением согласованности в качественном поведении решений дифференциальной и соответствующей разностной задач, возникают в случае $\mathrm{Re}\lambda_i < 0$, особенно при большом разбросе собственных значений λ_i, $i = 1, 2, \ldots, n$. Рассмотрим для простоты лишь случай вещественного спектра. Тогда выбор множителей s_i, $i = 1, 2, \ldots, n$, вида (11), как и в [5], естественно подчинить ограничениям $0<s_i<1$, $i=1, 2, \dots, n$, или очевидным более строгим требованиям $\exp(-\tau\|A\|)\le s_i<1$, $i=1,2,\dots, n$. Легко проверить непосредственно, что при $-\|A\|\le\mu<0$ выполнение, например, последних условий может быть обеспечено для любого $\tau>0$, если корректор σ выбрать по правилу:

$$\sigma=\exp(\mu\tau)\rho[-\tau(\|A\|+\mu)]/\rho(\mu\tau). \quad (12)$$

В силу (12), если $\tau\to0$, то $\sigma\to1$, при этом (см. (11)) и $s_i\to1$, $i=1,2,\dots,$ n. И аналогично, $\sigma\to0$ и $s_i\to0$ ($i=1,2,\dots, n$), когда $\tau\to\infty$. При этом, очевидно $0<\sigma<1$ для любого фиксированного $\tau>0$. Заметим далее, что в (10) вектор $g-SA^{-1}b$ отличается от истинного положения равновесия системы (2) слагаемым $\tau\rho(\mu\tau)(1-\sigma)(a-\mu A^{-1}b)$, которое зануляется при $\sigma=1$ для любых μ и a, а при $\sigma\neq1$ — лишь в случае

$$Aa=\mu b. \quad (13)$$

Для однородной системы ($b=0$) выполнение условия (13) может быть обеспечено без обращения матрицы A выбором $a=0$ (при любом μ). Вычисляя в этом случае, как и в [7], значения μ через отношение Релея на приближенном решении, мы не только обеспечим свойство точности метода на гармониках системы, но и гарантируем для случая симметричной матрицы A монотонное поведение этого решения вдоль разностных траекторий, приближающих, согласно данному методу, отличные от гармоник нетривиальные решения. Последнее легко проверить непосредственно, если учесть, что выбор σ по правилу (12) гарантирует выполнение естественного условия $s_j<s_k$ для $\lambda_j<\lambda_k$.

Так как в общем случае неоднородной системы (2) производная $u'(t)$ является решением соответствующей (2) однородной системы, а для однородного случая выбор значений всех параметров был выше обсужден, то тем самым принципиально описан вычислительный алгоритм для приближенного нахождения значений $u'(t)$, которые могут быть уже использованы при конструировании соответствующих (5) численных методов для случая исходной неоднородной системы.

Как мы уже отмечали, $s_i\to0$ ($i=1,2,\dots$ n) при $\tau\to\infty$. Поэтому в случае $b\neq0$ при выборе способа задания a ориентироваться на выполнение условия (13) естественно лишь в тех случаях, когда построенный метод будет точным, например, на решениях типа

$$u(t)=u^i(t)=-A^{-1}b+C_i\xi^i\exp(\lambda_i t).$$

Поскольку для решений такого вида справедливо соотношение

$$u(t)-u'(t)/\lambda_i=-A^{-1}b,$$

то выбор в качестве a вместо $\mu A^{-1}b$ вектора, близкого к $u'-\lambda_i u$, становится естественным, по крайней мере, для систем, моделирующих процессы с наличием регулярного режима [1]. Учитывая простую покоординатную связь посредством (5) векторов a и $\hat{y}$, выбор a можно связать и с неизвестным значением $\hat{y}$. Это в сочетании с описанным выше алгоритмом для приближенного нахождения значений производной $u'(t)$ делает вполне понятным, например, следующий выбор вычислительной процедуры:

$$v^0=Ay^0+b,\ \hat{v}=v+\tau\rho(\mu\tau)[(1-\sigma)\mu I+\sigma A]v,$$

$$\hat{y}=y+\tau\rho(\mu\tau)[(1-\sigma)\hat{v}+\sigma(Ay+b)]/[\exp(\mu\tau)-\tau\rho(\mu\tau)\mu\sigma],$$

$$\mu=\mu_0(v),\ \mu_0(x)=(Ax,x)/(x,x),\ \rho(z)=(\exp z-1)/z,$$

$$\sigma=\exp(\mu\tau)\rho[-\tau(\|A\|+\mu)]/\rho(\mu\tau).$$

Здесь через y^0 обозначено начальное состояние u^0 наблюдаемой траектории системы (2).

В заключение отметим, что предлагаемые расчетные формулы можно интерпретировать по аналогии с классическими разностными схемами с весами (см., напр., [1]), однако весовые коэффициенты здесь, в отличие от классического случая, не фиксированы, а изменяются во времени, адаптируясь к численно наблюдаемой траектории.

Список литературы

1. Самарский А. А. Теория разностных схем. М., 1983.
2. Деккер К., Вервер Я. Устойчивость методов Рунге—Кутты для жестких нелинейных дифференциальных уравнений. М., 1988.
3. Бобков В. В., Мандрик П. А. // Дифференц. уравнения. 1989. Т. 25. № 7. С. 1171.
4. Бобков В. В. // Вестн. Белорус. ун-та. Сер. 1: Физ. Мат. Мех. 1987. № 2. С. 72.
5. Бобков В. В., Бобкова Н. А. // Там же. 1991. № 1. С. 57.
6. Бобков В. В. // Дифференц. уравнения 1983. Т. 19. № 7. С. 1115.
7. Бобков В. В. // Там же. 1985. Т. 21. № 7. С. 1117.

Поступила в редакцию 02.12.91.

УДК 517.518

В. И. КОРЗЮК

ОБ ОПЕРАТОРАХ ОСРЕДНЕНИЯ С ПЕРЕМЕННЫМ ШАГОМ. II

Работа посвящена дальнейшему изучению свойств операторов осреднения В. И. Буренкова [1, 2] с целью последующего их использования при доказательстве разрешимости граничных задач дифференциальных уравнений. Будем пользоваться обозначениями и понятиями [3] без предварительного объяснения, а также продолжим нумерацию формул, свойств и лемм. Поэтому при ссылках на меньшую нумерацию необходимо смотреть работу [3].

Обобщим утверждение J-6 на случай производных любого конечного порядка α. Пусть функция $a(x)\in C^{|\alpha|}(\bar\Omega)$. Для $u\in H^{|\alpha|}(\Omega)$

$$F_\alpha^1 u(x) = D^\alpha[aJ_k u - J_k(au)] =$$

$$= \sum_{m=0}^{\infty}\sum_{\beta\le\alpha}\sum_{\gamma\le\alpha-\beta}\sum_{\eta\le\alpha-\beta-\gamma} a_\beta a_\gamma a_\eta D^\beta\psi_m(x)\times \tag{13}$$

$$\times(-1)^{|\alpha-\beta-\gamma-\eta|}\frac{1}{\delta_{mk}^n}\int_\Omega D_y^{\alpha-\beta-\gamma-\eta}\left\{\omega\left(\frac{x-y}{\delta_{mk}}\right)D_x^\gamma D_y^\eta[a(x)-a(y)]\right\}u(y)dy,$$

где a_β, a_γ, a_η—коэффициенты формулы Лейбница производной произведения двух функций, $D_x = \left(\frac{\partial}{\partial x_1},\ldots,\frac{\partial}{\partial x_n}\right)$, $D_y\left(\frac{\partial}{\partial y_1},\ldots,\frac{\partial}{\partial y_n}\right)$. Аналогично,

$$F_\alpha^{(2)}u(x) = D^\alpha[aJ_k^* u - J_k^*(au)] = \tag{14}$$

$$= \sum_{m=0}^{\infty}\sum_{\gamma\le\alpha}\sum_{\beta\le\alpha-\beta}\sum_{\eta\le\beta} a_\beta a_\gamma a_\eta(-1)^{|\alpha-\beta-\eta|}\frac{1}{\delta_{mk}^n}\times$$

$$\times\int_\Omega D_y^{\alpha-\beta-\gamma}\left\{\omega\left(\frac{x-y}{\delta_{mk}}\right)D^{\beta-\eta}\psi_m(y)D_y^\eta[a(x)-a(y)]\right\}u(y)dy.$$

Отметим, что в (13) и (14) $D_x^\gamma\, D_y^\eta[a(x)-a(y)]=0$ при $\gamma\ne0$ и $\eta\ne0$. В силу симметрии при $\eta=0$ и некотором $\gamma\le\alpha$ найдется слагаемое, которое вместо множителя в подинтегральном выражении $D_x^\gamma[a(x)-a(y)]$ будет содержать $D_y^\gamma[a(x)-a(y)]$. Поэтому (13) и (14) представляют собой сумму слагаемых, в подинтегральное выражение которых будут входить коэффициенты $D_x^\gamma a(x) - D_y^\gamma a(y)$ для всех $\gamma\le\alpha$. На основании сказанного, используя схему доказательства J—6, примененного к выражениям (13) и (14), получим следующее утверждение

J—7. Для любых $u\in H^{|\alpha|}(\Omega)$, $a\in C^{|\alpha|}(\bar\Omega)$ и $k\in N$ можно выбрать так числа $\delta_{mk}=\delta_{mk}(u)$, что

$$\|F_\alpha^{(i)}u\|_{L_2(\Omega)}\le\frac{1}{k}\|u\|_{H^{|\alpha|-1}(\Omega)},\quad i=1,2. \tag{15}$$

J—8. Для любых $u\in L_2(\Omega)$ можно выбрать так числа $\delta_{mk}=\delta_{mk}(u)$, что