EESTI NSV TEADUSTE AKADEEMIA TOIMETISED. XII KÖIDE
FÜÜSIKA-MATEMAATIKA- JA TEHNIKATEADUSTE SEERIA. 1963, NR. 4

ИЗВЕСТИЯ АКАДЕМИИ НАУК ЭСТОНСКОЙ ССР. ТОМ XII
СЕРИЯ ФИЗИКО-МАТЕМАТИЧЕСКИХ И ТЕХНИЧЕСКИХ НАУК. 1963, № 4

ОБ ОЦЕНКЕ ПОГРЕШНОСТИ МЕТОДА КОНЕЧНЫХ РАЗНОСТЕЙ ПРИ РЕШЕНИИ КРАЕВЫХ ЗАДАЧ ДЛЯ ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ ПРОИЗВОЛЬНОГО ПОРЯДКА

Р. ЮРГЕНСОН

Для изучения погрешности метода конечных разностей на практике часто применяют метод, предложенный Л. Коллацом [3]. По этому методу погрешности определяются решением системы уравнений для погрешностей, в которой высшие производные заменены их максимальными значениями.

Такой метод строг в случае определенных уравнений второго порядка. При уравнениях более высокого порядка оценки погрешности, в общем, не являются строгими, а вычисления, необходимые для оценки погрешности, — довольно трудоемки, так как, кроме нахождения оценок для высших производных, придется решать систему уравнений нередко с большим числом неизвестных.

Целью настоящей статьи является вывести строгие оценки погрешности, которые были бы применимы для различных разностных методов приближенного решения краевых задач для дифференциальных уравнений произвольного порядка.

Наши исследования основаны на использовании матрицы, в некотором смысле близкой к функции Грина (см., напр., [4]). Эту матрицу мы будем называть дискретной функцией Грина. Ее точное определение дается во втором пункте.

1. Система уравнений для погрешностей

Рассмотрим дифференциальное уравнение

$$x^{(n)} + \sum_{k=0}^{n-1} p_k(s)\, x^{(k)} = f(s) \tag{1}$$

с краевыми или начальными условиями

$$U_j(x) \equiv \sum_{k=0}^{n-1} [a_{jk} x^{(k)}(a) + b_{jk} x^{(k)}(b)] = c_j \tag{2}$$

$$(j = 1, 2, \ldots, n),$$

где $p_k(s)$, $f(s)$ — непрерывные на отрезке $a \leqslant s \leqslant b$ функции.

Предположим, что краевая задача $\{(1), (2)\}$ имеет единственное решение.

Для приближенного решения задачи $\{(1), (2)\}$ разобьем отрезок $a \leqslant s \leqslant b$ на m равных частей точками $s_i = a + ih$, где $h = \frac{b-a}{m}$. Дифференциальное уравнение (1) рассмотрим в точках s_i и заменим в нем производные (в точках s_i) разностными соотношениями, пользуясь формулами численного дифференцирования

$$x^{(k)}(s_i) = \frac{D_i^{(k)}(x(s_\nu))}{h^k} - R_i^{(k)} \qquad (k = 0, 1, \ldots, n). \tag{3}$$

Также заменим разностными соотношениями производные $x^{(k)}(a)$ и $x^{(k)}(b)$ в краевых условиях (2)

$$x^{(k)}(a) = \frac{D_a^{(k)}(x(s_\nu))}{h^k} - R_a^{(k)} \qquad (k = 0, 1, \ldots, n-1), \tag{4}$$

$$x^{(k)}(b) = \frac{D_b^{(k)}(x(s_\nu))}{h^k} - R_b^{(k)} \qquad (k = 0, 1, \ldots, n-1). \tag{5}$$

Здесь

$$\frac{D_i^{(k)}(x(s_\nu))}{h^k}, \qquad \frac{D_a^{(k)}(x(s_\nu))}{h^k}, \qquad \frac{D_b^{(k)}(x(s_\nu))}{h^k}$$

— некоторые линейные выражения относительно $x(s_\nu)$, аппроксимирующие производные $x^{(k)}(s)$ в точках s_i, a и b соответственно, а $R_i^{(k)}$, $R_a^{(k)}$, $R_b^{(k)}$ — остаточные члены соответствующих формул численного дифференцирования.

При этом $D_i^{(0)}(x(s_\nu)) = x_i$, $\quad D_a^{(0)}(x(s_\nu)) = x_0 = a$,

$$D_b^{(0)}(x(s_\nu)) = x_m = b, \quad R_i^{(0)} = R_a^{(0)} = R_b^{(0)} = 0.$$

После описанных подстановок получается следующая система линейных алгебраических уравнений

$$\frac{D_i^{(n)}(x(s_\nu))}{h^n} + \sum_{k=0}^{n-1} p_{ki} \frac{D_i^{(k)}(x(s_\nu))}{h^k} = f_i + R_i \qquad (i = \tilde{p}, \tilde{p}+1, \ldots, \tilde{q}), \tag{6}$$

$$V_j(x(s_\nu)) \equiv \sum_{k=0}^{n-1} \left[a_{jk} \frac{D_a^{(k)}(x(s_\nu))}{h^k} + b_{jk} \frac{D_b^{(k)}(x(s_\nu))}{h^k} \right] = c_j + R_j^{(\text{кр})} \qquad (j = 1, 2, \ldots, n), \tag{7}$$

где

$$p_{ki} = p_k(s_i), \quad f_i = f(s_i), \quad R_i = R_i^{(n)} + \sum_{k=1}^{n-1} p_{ki} R_i^{(k)} \qquad (i = \tilde{p}, \tilde{p}+1, \ldots, \tilde{q}),$$

$$R_j^{(\text{кр})} = \sum_{k=1}^{n-1} [a_{jk} R_a^{(k)} + b_{jk} R_b^{(k)}] \qquad (j = 1, 2, \ldots, n).$$

Натуральные числа $\tilde{p}$, $\tilde{q}$ и функции $D_i^{(k)}$, $D_a^{(k)}$, $D_b^{(k)}$ нужно выбирать таким образом, чтобы число неизвестных и число уравнений в системе $\{(6), (7)\}$ совпадали. Пусть в систему $\{(6), (7)\}$ входят неизвестные* $x(s_p)$, $x(s_{p+1}), \ldots, x(s_q)$. Тогда должно иметь место равенство

* Отметим, что некоторые s_j могут выйти за пределы промежутка $a \leqslant s \leqslant b$.

$$q - p + 1 = n + \tilde{q} - \tilde{p} + 1. \tag{8}$$

Решениями системы $\{(6), (7)\}$ являются значения $x(s_i)$ точного решения $x(s)$ задачи $\{(1), (2)\}$ в точках s_i.

Систему уравнений метода конечных разностей получим, отбросив в системе $\{(6), (7)\}$ члены R_i $(i = \tilde{p}, \tilde{p}+1, \ldots, \tilde{q})$, и $R_j^{(\text{кр})}$ $(j = 1, 2, \ldots, n)$

$$\frac{D_i^{(n)}(x_\nu)}{h^n} + \sum_{k=0}^{n-1} p_{ki} \frac{D_i^{(k)}(x_\nu)}{h^k} = f_i \qquad (i = \tilde{p}, \tilde{p}+1, \ldots, \tilde{q}), \tag{9}$$

$$V_j(x_\nu) = c_j \qquad (j = 1, 2, \ldots, n). \tag{10}$$

При практическом составлении уравнений (9) для конкретных задач во всех точках, кроме некоторых крайних, используют одни и те же формулы численного дифференцирования. Это значит, что почти во всех точках функции $D_i^{(k)}$ $(k = 1, 2, \ldots, n)$, не зависят от i. Пусть $D_i^{(k)}$ не зависят от i при $i = \bar{p}, \bar{p}+1, \ldots, \bar{q}$ $(\bar{p} \geqslant \tilde{p},\ \bar{q} \leqslant \tilde{q})$. Для этих i обозначим

$$D_i^{(n)}(x_\nu) = \sum_{j=-v_1}^{v_2} d_j x_{i+j} \qquad (i = \bar{p}, \bar{p}+1, \ldots, \bar{q}), \tag{11}$$

где постоянные d_j, v_1, v_2 определены соответствующей формулой численного дифференцирования.

Предположим, что выражения (11) содержат при $i = \bar{p}$ и $i = \bar{q}$ соответственно крайние неизвестные x_p и x_q.

Тогда имеют место равенства

$$\bar{p} - v_1 = p, \tag{12}$$

$$\bar{q} + v_2 = q. \tag{13}$$

Нашей задачей является нахождение оценок для $\varepsilon_i = x(s_i) - x_i$. Для получения таких оценок составим систему уравнений для погрешностей, вычитая из системы $\{(6), (7)\}$ систему $\{(9), (10)\}$

$$\frac{D_i^{(n)}(\varepsilon_\nu)}{h^n} + \sum_{k=0}^{n-1} p_{ki} \frac{D_i^{(k)}(\varepsilon_\nu)}{h^k} = R_i \qquad (i = \tilde{p}, \tilde{p}+1, \ldots, \tilde{q}), \tag{14}$$

$$V_j(\varepsilon_\nu) = R_j^{(\text{кр})} \qquad (j = 1, 2, \ldots, n). \tag{15}$$

Систему $\{(14), (15)\}$ преобразуем в систему с однородными краевыми условиями. Для этого сделаем в упомянутой системе замену переменной

$$\varepsilon_i = u_i + t_i \qquad (i = p, p+1, \ldots, q), \tag{16}$$

где t_i удовлетворяет условиям

$$\sum_{j=-v_1}^{v_2} d_j t_{i+j} = 0 \qquad (i = \bar{p}, \bar{p}+1, \ldots, \bar{q}), \tag{17}$$

$$\left.\begin{aligned} &\frac{D_i^{(n)}(t_\nu)}{h^n}+\sum_{k=0}^{n-1} p_{ki}\,\frac{D_i^{(k)}(t_\nu)}{h^k}=R_i \\ &(i=\tilde{p},\ \tilde{p}+1,\ \ldots,\ \bar{p}-1,\ \bar{q}+1,\ \ldots,\ \tilde{q}), \\ &V_j(t_\nu)=R_j^{(\text{кр})} \qquad (j=1,2,\ldots,n). \end{aligned}\right\} \tag{18}$$

При практическом определении элементов t_i пишем их в виде

$$t_i=e_0z_0(i)+e_1z_1(i)+\ldots+e_{v_2+v_1-1}\,z_{v_2+v_1-1}\,(i),$$

где $z_j(i)$ $(j=0,\ 1,\ldots,\ v_2+v_1-1)$ являются линейно-независимыми решениями* уравнений (17).

Докажем, что число условий (18) равно числу постоянных e_j, т. е. равно v_2+v_1.

Число уравнений (18) равно $n+\tilde{q}-\tilde{p}-(\bar{q}-\bar{p})$.

Вычтем из равенства (13) равенство (12)

$$v_2+v_1+\bar{q}-\bar{p}=q-p.$$

Отсюда

$$v_2+v_1=q-p-(\bar{q}-\bar{p}),$$

или, учитывая равенство (8),

$$v_2+v_1=n+\tilde{q}-\tilde{p}-(\bar{q}-\bar{p}),$$

что и требовалось доказать.

Итак, для определения постоянных e_j получим из системы условий (18) систему уравнений, в которой число неизвестных и число уравнений совпадают. Предположим, что определитель этой системы отличен от нуля.

Вследствие замены переменной (16) система уравнений $\{(14),\ (15)\}$ заменяется следующей системой

$$\frac{\sum\limits_{j=-v_1}^{v_2} d_j u_{i+j}}{h^n}+\sum_{k=0}^{n-1} p_{ki}\,\frac{D_i^{(k)}(u_\nu)}{h^k}=R_i-\sum_{k=0}^{n-1} p_{ki}\,\frac{D_i^{(k)}(t_\nu)}{h^k} \qquad (i=\bar{p},\ \bar{p}+1,\ldots,\ \bar{q}), \tag{19}$$

$$\left.\begin{aligned} &\frac{D_i^{(n)}(u_\nu)}{h^n}+\sum_{k=0}^{n-1} p_{ki}\,\frac{D_i^{(k)}(u_\nu)}{h^k}=0 \\ &(i=\tilde{p},\ \tilde{p}+1,\ldots,\ \bar{p}-1,\ \bar{q}+1,\ldots,\ \tilde{q}) \\ &V_j(u_\nu)=0 \qquad (j=1,\ 2,\ldots,\ n). \end{aligned}\right\} \tag{20}$$

С помощью дискретной функции Грина эта система преобразуется к виду более подходящему для оценки погрешности.

---

* Методы решения разностных уравнений такого типа изложены, например, в монографии А. О. Гельфонда [1].

2. Дискретная функция Грина и ее основное свойство

Введем понятие дискретной функции Грина.

Определение. Дискретной функцией Грина системы

$$\sum_{j=-v_1}^{v_2} d_j u_{i+j} = 0 \qquad (i = \bar{p}, \bar{p}+1, \ldots, \bar{q}) \tag{21}$$

при краевых условиях (20) называется $(\bar{q}-\bar{p}+1)\times(q-p+1)$-матрица G, элементы g_{ik} которой (как функции от первого индекса) удовлетворяют условиям

$$\sum_{j=-v_1}^{v_2} d_j g_{i+j,\,k} = \begin{cases} 0 & \text{при } i \neq k, \\ 1 & \text{при } i = k \end{cases} \tag{22}$$

$$(i, k = \bar{p}, \bar{p}+1, \ldots, \bar{q}),$$

$$\left.\begin{array}{l} \dfrac{D_i^{(n)}(g_{\nu k})}{h^n} + \displaystyle\sum_{j=0}^{n-1} p_{ji} \dfrac{D_i^{(j)}(g_{\nu k})}{h^j} = 0 \\ (i = \tilde{p}, \tilde{p}+1, \ldots, \bar{p}-1, \bar{q}+1, \ldots, \tilde{q};\ k = \bar{p}, \bar{p}+1, \ldots, \bar{q}) \\ V_j(g_{\nu k}) = 0 \qquad (j = 1, 2, \ldots, n;\ k = \bar{p}, \bar{p}+1, \ldots, \bar{q}). \end{array}\right\} \tag{23}$$

Для каждого фиксированного k $(k = \bar{p}, \bar{p}+1, \ldots, \bar{q})$ система $\{(22), (23)\}$ состоит из $n + \tilde{q} - \tilde{p} + 1 = q - p + 1$ условий. Если определитель этой системы отличается от нуля, то существует единственная дискретная функция Грина задачи $\{(21), (20)\}$.

Основное свойство дискретной функции Грина определяется следующей леммой.

Лемма. *Решение u_i системы*

$$\frac{1}{h^n} \sum_{j=-v_1}^{v_2} d_j u_{i+j} = r_i \qquad (i = \bar{p}, \bar{p}+1, \ldots, \bar{q}), \tag{24}$$

где r_i — произвольные постоянные, при условиях (20), *выражается формулой*

$$u_i = h^n \sum_{k=\bar{p}}^{\bar{q}} g_{ik} r_k \qquad (i = p, p+1, \ldots, q). \tag{25}$$

Доказательство. В силу (23) элементы u_i удовлетворяют условиям (20).

Докажем, что u_i, заданные формулой (25), превращают систему (24) в тождество.

Имеем

$$\frac{1}{h^n} \sum_{j=-v_1}^{v_2} d_j u_{i+j} = \frac{1}{h^n} \sum_{j=-v_1}^{v_2} d_j h^n \sum_{k=\bar{p}}^{\bar{q}} g_{i+j,\,k} r_k = \sum_{k=\bar{p}}^{\bar{q}} \sum_{j=-v_1}^{v_2} d_j g_{i+j,\,k} r_k.$$

В силу (22) последнее выражение равно r_i, что и требовалось доказать.

Для практического отыскания матрицы G элементы g_{ik} целесообразно писать в виде

$$g_{ik} = \begin{cases} a_0 z_0(i) + a_1 z_1(i) + \ldots + a_{v_2+v_1-1} z_{v_2+v_1-1}(i) & \text{при } i \leqslant k, \\ b_0 z_0(i) + b_1 z_1(i) + \ldots + b_{v_2+v_1-1} z_{v_2+v_1-1}(i) & \text{при } i \geqslant k \end{cases}$$

$$(i = p, p+1, \ldots, q; \; k = \bar{p}, \bar{p}+1, \ldots, \bar{q}),$$

где $z_j(i)$ $(j = 0, 1, \ldots, v_2 + v_1 - 1)$ — линейно-независимые решения системы (21).

При таком построении элементов g_{ik} условия (22) будут выполнены для $i = \bar{p}, \bar{p}+1, \ldots, k - v_2, k + v_1, \ldots, \bar{q}$.

Постоянные a_j, b_j $(j = 0, 1, \ldots, v_2 + v_1 - 1)$ определяются из условий (22) (при $i = k - v_2 + 1, k - v_2 + 2, \ldots, k + v_1 - 1$) и (23). На основании равенств (8), (12), (13) легко доказать, что число этих условий равно $2(v_2 + v_1 - 1)$.

Отметим, что идея введения понятия дискретной функции Грина возникла при изучении работ И. Шредера [7, 8], в которых элементы g_{ik} найдены для задачи

$$u_{i+1} - 2u_i + u_{i-1} = 0 \qquad (i = 1, 2, \ldots, m-1),$$

$$a_0 u_0 - a_1 l \, \frac{u_1 - u_0}{h} = 0,$$

$$b_0 u_m + b_1 l \, \frac{u_m - u_{m-1}}{h} = 0$$

(l — длина отрезка, на котором рассматривается дифференциальное уравнение; $a_0, a_1, b_0, b_1 \geqslant 0$).

Для некоторых систем, являющихся системами метода конечных разностей при решении дифференциальных уравнений второго порядка, дискретные функции Грина использованы и другими авторами [5, 6].

3. Оценка погрешности

Сделаем в системе $\{(19), (20)\}$ замену переменной

$$\sum_{j=-v_1}^{v_2} d_j u_{i+j} = z_j \qquad (i = \bar{p}, \bar{p}+1, \ldots, \bar{q}). \tag{26}$$

Тогда при краевых условиях (20), в силу леммы (из пункта 2),

$$u_l = h^n \sum_{k=\bar{p}}^{\bar{q}} g_{ik} z_k, \tag{27}$$

и система $\{(19), (20)\}$ принимает вид

$$z_i + \sum_{k=0}^{n-1} p_{ki} \frac{1}{h^k} D_i^{(k)} \left(h^n \sum_{j=\overline{p}}^{\overline{q}} g_{\nu j} z_j\right) = R_i - \sum_{k=0}^{n-1} p_{ki} \frac{D_i^{(k)}(t_\nu)}{h^k}$$

$$(i = \overline{p}, \overline{p}+1, \ldots, \overline{q}),$$

или, в силу линейности функций $D_i^{(k)}$, вид

$$z = Az + R, \tag{28}$$

где

$$z = (z_i), \quad R = \left(R_i - \sum_{k=0}^{n-1} p_{ki} \frac{D_i^{(k)}(t_\nu)}{h^k}\right),$$

$$Az = \left(-\sum_{k=0}^{n-1} p_{ki} h^{n-k} \sum_{j=\overline{p}}^{\overline{q}} D_i^{(k)}(g_{\nu j} z_j)\right).$$

Эту систему рассмотрим как операторное уравнение в пространстве $(\overline{q} - \overline{p} + 1)$-мерных векторов $R_{\overline{q}-\overline{p}+1}$ и введем норму

$$\|z\| = \max_{i=\overline{p}, \overline{p}+1, \ldots, \overline{q}} |z_i|.$$

Обозначим оценки

$$\max_{i=\overline{p}, \overline{p}+1, \ldots, \overline{q}} \left| R_i - \sum_{k=0}^{n-1} p_{ki} \frac{D_i^{(k)}(t_\nu)}{h^k} \right| \leqslant R^*,$$

$$|t_i| \leqslant \tau_i \quad (i = \overline{p}, \overline{p}+1, \ldots, \overline{q}), \qquad \max_{i=\overline{p}, \overline{p}+1, \ldots, \overline{q}} |p_{ki}| \leqslant \pi_k,$$

$$\max_{i=\overline{p}, \overline{p}+1, \ldots, \overline{q}} h^{n-k} \sum_{k=\overline{p}}^{\overline{q}} |D_i^{(k)}(g_{\nu j})| \leqslant \mu_k \qquad (k = 0, 1, \ldots, n-1).$$

Легко доказать, что имеют место оценки норм

$$\|R\| \leqslant R^*, \qquad \|A\| \leqslant \mu = \sum_{k=0}^{n-1} \pi_k \mu_k.$$

Предположим, что $\mu < 1$. Тогда на основании теоремы Банаха (см., напр., [2], глава V, § 2) система (28) имеет единственное решение $\{z_i\}$ $(i = \overline{p}, \overline{p}+1, \ldots, \overline{q})$, причем справедливо неравенство

$$\max_{i=\overline{p}, \overline{p}+1, \ldots, \overline{q}} |z_i| \leqslant \frac{R^*}{1-\mu}.$$

Оценки для погрешностей ε_i получим из этого неравенства с помощью соотношений (16) и (27)

$$|\varepsilon_i| \leqslant \tau_i + h^n \sum_{k=\overline{p}}^{\overline{q}} |g_{ik}| \frac{R^*}{1-\mu} \; (i = p, p+1, \ldots, q). \tag{29}$$

Правая часть последнего неравенства зависит от дискретной функции Грина и от остаточных членов формул численного дифференцирования $R_i^{(k)}$, $R_a^{(k)}$, $R_b^{(k)}$. Члены $R_i^{(k)}$, $R_a^{(k)}$, $R_b^{(k)}$, в свою очередь, зависят от высших производных решения задачи $\{(1), (2)\}$. Следовательно, точность оценки (29) существенно зависит от того, насколько высшие производные близки к постоянным.

Функция G и элементы t_i определяются так, как указано в пп. 1 и 2. Для удобства практического пользования оценкой (29) можно составить таблицы, в которых даются G, t_i и оценки μ_k в случае конкретных разностных систем.

ЛИТЕРАТУРА

1. А. О. Гельфонд, Исчисление конечных разностей. М., 1959.
2. Л. В. Канторович, Г. П. Акилов, Функциональный анализ в нормированных пространствах. М., 1959.
3. Л. Коллац, Численные методы решения дифференциальных уравнений. М., 1953.
4. М. А. Наймарк, Линейные дифференциальные операторы. М., 1954.
5. А. Н. Тихонов, А. А. Самарский, Об однородных разностных схемах, Ж. вычисл. матем. и матем. физ., **1**, 1, 5—63, 1961.
6. M. A. Abdel-Messih, A Green's Function Analogue for Ordinary Linear Difference Equations, Proc. Math. and Phys. Soc. Egypt, **22**, 43—51, 1958.
7. J. Schröder, Über die Differenzenverfahren bei nichtlinearen Randwertaufgaben I, Z. angew. Math. und Mech., **36**, 9—10, 319—331, 1956.
8. J. Schröder, Über die Differenzenverfahren bei nichtlinearen Randwertaufgaben II, Z. angew. Math. und Mech., **36**, 11—12, 443—455, 1956.

Институт физики и астрономии
Академии наук Эстонской ССР

Поступила в редакцию
18. X 1963

DIFERENTSIMEETODI VEAHINNANGUST MISTAHES JÄRKU DIFERENTSIAALVÕRRANDITE RAJAÜLESANNETE LAHENDAMISEL

R. Jürgenson

Resümee

Veahinnangute aluseks on punktis 2 defineeritav diskreetne Greeni funktsioon. Viimase abil teisendatakse vigade võrrandisüsteem (süsteem $\{(14), (15)\}$) kujule, kus veahinnangud on kergesti leitavad.

Punktis 3 toodud ranged veahinnangud on rakendatavad enamiku praktikas kasutatavate diferentsimeetodite korral.

Eesti NSV Teaduste Akadeemia
Füüsika ja Astronoomia Instituut

Saabus toimetusse
18. X 1963

ON THE ERROR ESTIMATION OF THE FINITE DIFFERENCE METHOD IN SOLVING THE BOUNDARY VALUE PROBLEM OF THE DIFFERENCE EQUATION OF ARBITRARY ORDER

R. Jürgenson

Summary

The error estimation is based on the discrete Green's function, defined in section 2. This function has been used to transform the system of error equations (section 1) into a form in which the error estimation can be easily carried out.

The strict error estimations given in section 3 are applicable to the majority of the finite difference methods used in practice.

Academy of Sciences of the Estonian S.S.R.,
Institute of Physics and Astronomy

Received
Oct. 18th, 1963