On the error bounds for the ordinary differential equations of the first order

Shoichi Seino
Saburo Minato

1. Introduction

When we deal with a differential equation of any type, it frequently happens that its solution is not expressible in closed form, and even if a given differential equation possesses a closed form solution in terms of known functions, it also happens that the solution is represented by little tabulated functions.

In these cases, a numerical approach is the most convenient way for solving the differential equation. Many methods are available for solving it numerically, so it is important for us to choose the best one among them. An index of it is the error, for the error analysis gives a qualitative basis for an estimation of the computation value and it serves to give the guarantee of accuracy for a computed solution. When we get some approximate solution of the differential equation, we wish to know something about the size of the error. But the estimation of the error is so complicated that it is hard to get a practical and efficient estimate.

W. Uhlmann in Hamburg developes the error bound function, corresponding to the error, of the system of the ordinary differential equations of the first order, considering some kind of "majorant".

On the other hand, G. G. Dahlquist in Stockholm is working in the same line and he introduces the concept of majorant functions in his essential step.

We are interested in both methods, but unfortunately the publications referring to them are little, consequently numerical experiments on practical examples are needed before successful use of it. Especially on latter method, the algorithms are not yet completely specified, as he says.

We shall compare the Uhlmann's method with some rude one of the estimation of the error bounds of the above type of the differential equation.

For simplicity we shall ignore rounding error in obtaining the following error estimates, i.e., we estimate only the inherent or truncation errors. Rounding error will, of course, be propagated throughout the calculation, but we take the view that, knowing an upper bound for the inherent error from one of the estimate, we can calculate with a number of decimal sufficient to ensure that the rounding error remain within this bound. In principle it would not be difficult to derive similar estimates for the rounding errors.

2. Efficient error bound

We consider the initial value problems of a system of s ordinary differential equations of the first order

$$(2.1) \qquad y_i' = f_i(x, y_1(x), \cdots, y_s(x)), \qquad y_i(x_0) = y_{i0} \ (i = 1, 2, \cdots, s),$$

where the functions f_i possess $(r+1)$-th continuous partial derivatives with respect to $(s+1)$ variables.

We set $x_n = x_0 + nh$ (h is the step length, $n = 0, 1, \cdots, N+1$). In the following the existence of the solution will be assumed. We calculate approximate values y_{in} to the values $y_i(x_n)$ of the exact solutions at the points x_n by the way of the r-th order Adams interpolation method for $m+1 \leqq n \leqq N+1$ $(m \geqq r-1)$.

We abbreviate f_{in} to $f_i(x_n, y_{1n}, \cdots, y_{sn})$.

Let $P_{in}(x)$ be the polynomial of degree r such that $P_{in}(x_k) = f_{ik}$ $(k = n+1-r, \cdots, n+1)$, then

$$y_{i,n+1} = y_{in} + \int_{x_n}^{x_{n+1}} P_{in}(t)dt.$$

For $x_n \leqq x \leqq x_{n+1}$ $(n = m, \cdots, N)$ we define the functions $y_i^*(x)$;

$$(2.2) \qquad y_i^*(x) = y_{in} + \int_{x_n}^{x} P_{in}(t)dt,$$

and the defect $d_i(x)$:

$$d_i(x) = y_i^{*\prime}(x) - f_i(x, y_1^*(x), \cdots, y_s^*(x)).$$

From (2.2)

$$d_i(x) = P_{in}(x) - f_i(x, y_1^*(x), \cdots, y_s^*(x)), \quad \text{for } x_n \leqq x \leqq x_{n+1}.$$

We calculate the bound d_{iM} of $|d_i(x)|$.

Owing to the fact the functions $y_i^*(x)$ and $f_i(x, y_1^*(x), \cdots, y_s^*(x))$ are $(r+1)$-th continuously differentiable, there exist $(r+1)$-th continuously differentiable functions defined in the interval $[x_{n+1-r}, x_{n+2}]$ such that

$$g_i(x) = f_i(x, y_1^*(x), \cdots, y_s^*(x)) \qquad \text{for } x_n \leqq x \leqq x_{n+1}$$

and

$$g_i(x_k) = f_{ik} \qquad \text{for } k = n+1-r, \cdots, n+2.$$

By the well-known formula on interpolation polynomial,

$$|d_i(x)| \leqq M_{r+1} h^{r+1} \cdot \max_{x_{n+1-r} \leqq x \leqq x_{n+1}} |g_i^{(r+1)}(x)| \quad \text{for } x_n \leqq x \leqq x_{n+1} (n = m, \cdots, N),$$

where

$$M_r = \max_{r-1 \leqq t \leqq r} \left| \frac{(t-1)(t-2)\cdots\cdots(t-r)}{r!} \right|.$$

We define the error $\varepsilon_i(x)$ such that

$$\varepsilon_i(x) = y_i^*(x) - y_i(x),$$

and we put the initial error bound E_i : $|\varepsilon_i(x_m)| \leqq E_i$.

In addition to these, we define the Lipschitz quotient $L_{ik}(x)$ such that

(2.3)

$$L_{ik}(x) = \frac{f_i(x, y_1(x), \cdots, y_{k-1}(x), y_k^*(x), \cdots, y_s^*(x)) - f_i(x, y_1(x), \cdots, y_k(x), y_{k+1}^*(x), \cdots, y_s^*(x))}{y_k^*(x) - y_k(x)}.$$

We assume that the functions $L_{ik}(x)$ are continuous and bounded in $x_m \leqq x \leqq x_N$, i.e., $l_{ik} \leqq L_{ik}(x) \leqq L_{ik}$, $|L_{ik}(x)| \leqq K_{ik} = \mathrm{Max}\{|l_{ik}|, |L_{ik}|\}$.

From (2.1) and (2.3)

$$\varepsilon_i'(x) = \sum_{k=1}^{s} L_{ik}(x)\varepsilon_k(x) + d_i(x) \quad (i=1,2,\cdots,s).$$

Then we have the following

[THEOREM I]

If the functions $E_i(x)$ are solutions of the initial value problem

$$E_i'(x) = \sum_{\substack{k=1\\k\neq i}}^{s} K_{ik}E_k(x) + L_{ii}E_i(x) + d_{iM}, \quad E_i(x_m) = E_i, \tag{2.4}$$

then, for $x_m \leqq x \leqq x_N$,

$$|\varepsilon_i(x)| \leqq E_i(x) \qquad (i=1,2,\cdots,s).$$

PROOF. We put

$$\varepsilon_i(x) = e^{-\alpha(x-x_m)}u_i(x), \text{ for any } \alpha \text{ such that } \alpha \geqq \mathop{\mathrm{Max}}_i K_{ii}.$$

Then we have

$$u_i'(x) = \sum_{\substack{k=1\\k\neq i}}^{s} L_{ik}(x)u_k(x) + (L_{ii}(x)+\alpha)u_i(x) + e^{\alpha(x-x_m)}d_i(x),$$

$$u_i(x_m) = \varepsilon_i(x_m).$$

We consider for $x \geqq x_m$ and $i=1,2,\cdots,s$,

$$z_i'(x) = \sum_{\substack{k=1\\k\neq i}}^{s} K_{ik}z_k(x) + (L_{ii}+\alpha)z_i(x) + e^{\alpha(x-x_m)}d_{iM}. \tag{2.5}$$

Each solution of (2.5) is written in the form

$$z_i(x) = z_{io}(x) + \sum_{k=1}^{s} z_k(x_m)z_{ik}(x),$$

where $z_{io}(x)$ is a particular solution of the inhomogeneous equation, $z_{ik}(x)$ are solutions of the homogeneous equation, and $z_{ik}(x_m) = \delta_{ik}$ $(i=1,2,\cdots,s;\ k=0,1,\cdots,s)$. We wish to show that all the $z_{ik}(x) \geqq 0$. To this end we put

$$z_{ik}^{(0)}(x) = z_{ik}(x_m),$$

and

$$z_{ik}^{(p+1)}(x) = z_{ik}(x_m) + \int_{x_m}^{x} \{ \sum_{\substack{j=1\\j\neq i}}^{s} K_{ij}z_{jk}^{(p)}(t) + (L_{ii}+\alpha)z_{ik}^{(p)}(t) + \delta_{ik}e^{\alpha(t-x_m)}d_{iM}\}dt.$$

Then we get $\lim_{p\to\infty} z_{ik}^{(p)}(x) = z_{ik}(x)$ by means of the Picard-Lindelöff theorem.

By induction, we have $z_{ik}^{(p)}(x)\geqq 0$, and it follows that $z_{ik}(x)\geqq 0$.

We denote the solutions of (2.5) with conditions $u_{io}(x_m)=\varepsilon_i(x_m)$ and $U_i(x_m)=E_i$ by $u_{io}(x)$ and $U_i(x)$ respectively. In addition, let

$$u_{i,p+1}(x)=\varepsilon_i(x_m)+\int_{x_m}^{x}\{\sum_{\substack{k=1\\k\neq i}}^{s}L_{ik}(t)u_{kp}(t)+(L_{ii}(t)+\alpha)u_{ip}(t)+e^{\alpha(t-x_m)}d_i(t)\}dt$$

for $p=0,1,2,\cdots$ and $x_m\leqq x\leqq x_N$.

We easily verify, by induction, that

$$|u_{ip}(x)|\leqq U_i(x)\qquad (i=1,2,\cdots,s;\ p=0,1,2,\cdots).$$

Again applying the Picard-Lindelöff theorem,

$$\lim_{p\to\infty} u_{ip}(x)=u_i(x),\ \text{and also } |u_i(x)|\leqq U_i(x).$$

We put

$$E_i(x)=e^{-\alpha(x-x_m)}U_i(x),$$

then $E_i(x)$ satisfies (2.4), and

$$|\varepsilon_i(x)|=|u_i(x)e^{-\alpha(x-x_m)}|\leqq U_i(x)\,e^{-\alpha(x-x_m)}=E_i(x).$$

Next result will be used later.

【COROLLARY】

For $x_m\leqq x\leqq x_N$,

$$|\varepsilon(x)|\leqq E(x)=\begin{cases} E\cdot e^{(x-x_m)L}+d_M\dfrac{e^{(x-x_m)L}-1}{L}, & \text{if } L\neq 0\\ E+d_M(x-x_m), & \text{if } L=0.\end{cases}$$

PROOF. Taking $i=1$ in (2.4), we can prove immediately.

3. Error bound for the improved polygon method

Hereafter we restrict our attention to the first order differential equation

$$y'=f(x,y),\quad y(x_o)=y_o. \tag{3.1}$$

We assume that the function $f(x,y)$

(a) is continuous in a domain D of the real (x,y) plane,

(b) is bounded as $|f(x,y)|\leqq M$ for any $(x,y)\in D$,

(c) satisfies the Lipschitz condition

$$|f(x,y)-f(x,y^*)|\leqq K|y-y^*|\quad \text{for any } (x,y) \text{ and } (x,y^*)\in D$$

(d) possesses bounded derivatives $|\frac{d^kf}{dx^k}|\leqq N_k$ $(k=1,2)$.

By y_n we denote the approximations of the values $y(x_n)$ of the exact solution of (3.1) at the points x_n $(n=1,2,\cdots)$.

Now we calculate the next y_{n+1} from y_n by the formula

$$y_{n+1}=y_n+hf_{n+\frac{1}{2}},$$

where $h=x_{n+1}-x_n$, $f_{n+\frac{1}{2}}=f(x_{n+\frac{1}{2}},y_{n+\frac{1}{2}})$, $x_{n+\frac{1}{2}}=x_n+\frac{1}{2}h$, and $y_{n+\frac{1}{2}}=y_n+\frac{1}{2}hf_n$.

The error in the approximate values y_n means the quantity

$$\varepsilon_n=y_n-y(x_n). \tag{3.2}$$

We put
$$\Delta\varepsilon_n=\varepsilon_{n+1}-\varepsilon_n$$
so that

$$\Delta\varepsilon_n=hf(x_{n+\frac{1}{2}},\ y_n+\frac{1}{2}hf_n)-\int_{x_n}^{x_{n+1}}F(x)dx$$

with $F(x)=f(x,y(x))$.

We set

$$\Delta\varepsilon_n=I_n+hJ_n, \tag{3.3}$$

where $I_n=hF(x_{n+\frac{1}{2}})-\int_{x_n}^{x_{n+1}}F(x)dx$, and $J_n=f(x_{n+\frac{1}{2}},y_{n+\frac{1}{2}}+\frac{1}{2}hf_n)-f(x_{n+\frac{1}{2}},y(x_{n+\frac{1}{2}}))$.

For the quadrature error I_n, the next well-known inequality holds:

$$|I_n|\leqq h^3N_2/24. \tag{3.4}$$

Using the Lipschitz condition (c), we obtain

$$|J_n|\leqq K|y_{n+\frac{1}{2}}+\frac{1}{2}hf_n-y(x_{n+\frac{1}{2}})|\leqq K\cdot((1+\frac{1}{2}hK)|\varepsilon_n|+\frac{1}{8}h^2N_1). \tag{3.5}$$

From (3.3), (3.4) and (3.5), it follows that

$$|\varepsilon_{n+1}|\leqq|\varepsilon_n|\ (1+hK+\frac{1}{2}h^2K^2)+\frac{1}{8}h^3(KN_1+\frac{1}{3}N_2).$$

Thus, we obtain the independent error bound from this recursive error estimate. The result obtained is stated in theorem form as follows.

【THEOREM II】

Let ε_n be the error defined by (3.2) *and E_n^* be the error bound of ε_n, then*

$$E_n^*=\frac{1}{8}h^2(N_1+\frac{N_2}{3K})\frac{(1+hK+\frac{1}{2}h^2K^2)^n-1}{1+\frac{1}{2}hK}.$$

4. Numerical examples

As the first example, we consider the inital value problem

$$y'=y-x,\quad y(0)=2.$$

We compute the values y_n which are the approximations of the exact solution $y(x)$ at the points x_n, using the Lunge-Kutta method of the fourth order with $h=0.1$. Table I shows the values of the error bounds corresponding to the methods in sec. 2 and 3.

Table I

n	x_n	y_n	d_M	$E(x_n)$	E_n^*
2	0.2	2.421 403	0.000 071	0.000 041	0.001 007
3	0.3	2.649 858	0.000 079	0.000 054	0.001 774
4	0.4	2.891 824	0.000 087	0:000 071	0.002 760
5	0.5	3.148 721	0.000 096	0.000 092	0.004 007
6	0.6	3.422 118	0.000 106	0.000 118	0.005 564
7	0.7	3.713 752	0.000 117	0.000 151	0.007 489
8	0.8	4.025 540	0.000 130	0.000 192	0.009 851
9	0.9	4.359 601	0.000 143	0.000 242	0.012 735
10	1.0	4.718 280	0.000 158	0.000 304	0.016 239

As another example, let's discuss the initial value problem

$$y' = 2xy, \qquad y(2) = 1.$$

The approximations y_n are calculated by the second order Adams interpolation method with the step length $h = 0.01$. We represent the result calculated for $E(x_n)$ aud E_n^* in Table II.

Table II

n	x_n	y_n	d_M	$E(x_n)$	E_n^*
5	2.05	1.224 461	0.000 074	0.000 004	0.000 019
10	2.10	1.506 820	0.000 098	0.000 010	0.000 055
15	2.15	1.863 587	0.000 130	0.000 019	0.000 122
20	2.20	2.316 377	0.000 174	0.000 034	0.000 240
25	2.25	2.893 613	0.000 234	0.000 056	0.000 452
30	2.30	3.632 813	0.000 315	0.000 088	0.000 829
35	2.35	4.583 710	0.000 425	0.000 135	0.001 499
40	2.40	5.812 498	0.000 577	0.000 205	0.002 698
45	2.45	7.407 642	0.000 787	0.000 307	0.004 860
50	2.50	9.487 869	0.001 077	0.000 455	0.008 789

5. Comment

These examples show that the value of the error bound $E(x)$ introduced by Uhlmann is smaller than E^*, and it increases slowly as n does. In comparing other methods it may be sure that the Uhlmann's method is very efficient. His procedure consists in taking the error bound function $E(x)$, corresponding to the error function $\varepsilon(x)$, which is the solution of the first order linear differential equation with constant coefficients. Generally speaking, it is not easy to seek bounds of Lipschitz quotients. Moreover, even if the function $E(x)$ is expressed in exact form, we must calculate the value of it numerically.

When the system of the differential equations is given, the numerical procedure of the error bound function $E(x)$ is more complicated.

References

1. W. Uhlmann: Fehlerabschätzungen bei Anfangswertaufgaben gewöhnlicher Differentialgleichungs-systeme 1. Ordnung, Z. angew. Math. Mech. Bd. 37 (1957).
2. W. Uhlmann: Zur Fehlerabschätzung bei Interpolationspolynomen, Z. angew. Math. Mech. Bd. 37 (1957).
3. L. Collatz: The Numerical Treatment of Differential Equations, Spriger-Verlag (1959).
4. G. G. Dahlquist: On Rigorous Error Bounds in the Numerical Solution of Ordinary Differential Equations, Proceeding of an Advanced Symposium Conducted by the Mathematics Research Center, ed. by D. Greenspan, John Willy & Sons, Inc. (1966).