Application of Extrapolation Methods to Initial Value problems in Ordinary Differential Equations

N. Oprea and D. Mărginean, ”Petru Maior” University of Tg. Mureş

Proceedings of the European Integration between Tradition and Modernity
”Petru Maior” University of Târgu-Mureş
Oktober, 22-23, 2009

Abstract

In this paper, using the polynomial extrapolation, we solve an initial value problem in ordinary differential equations. The aim of this paper is to compare with the fourth-order Runge-Kutta method on the basis of accuracy for a given number of function evaluations.

Keywords: polynomial extrapolation, superlinearly converges, basic steplength, the modified mid-point method, one-step and two-step method, weak stability, the existence of asymptotic expansions;

1 Polynomial extrapolation

In many situations in numerical analysis we wish to evaluate a number A_0, but we are able to compute only an approximation $A(h)$, where h is a positive parameter and where $A(h) \to A_0$ as $h \to 0$. Let us suppose that, for every fixed N, $A(h)$ possesses an asymptotic expansion of the form:

$$\begin{aligned} A(h) &= A_0 + A_1 h + A_2 h^2 + \ldots + A_N h^N + R_N(h), \\ R_N(h) &= O(h^{N+1}), \quad \text{as} \quad h \to 0, \end{aligned}$$

where the coefficients $A_0, A_1, \ldots, A_N$ are independent of h. We shall summarize this statement by writing:

$$A(h) \approx A_0 + A_1 h + A_2 h^2 + \ldots \tag{1}$$

Suppose that we have calculated $A(h_0)$ and $A(\frac{1}{2}\cdot h_0)$, so: $A(h_0) = A_0 + O(h_0)$ and $A(\frac{1}{2}h_0) = A_0 + O(h_0)$ as $h \to 0$. There exists a linear combination of $A(h_0)$ and $A(\frac{1}{2}h_0)$ which differs from A_0 by an $O(h_0^2)$ term:

$$2A\left(\frac{1}{2}h_0\right) - A(h_0) = A_0 - \frac{1}{2}A_2h_0^2 - \ldots = A_0 + O(h_0^2) \tag{2}$$

This is the basic idea of Richardson extrapolation. It can be extended in several ways: in addition to $A(h_0)$ and $A(\frac{1}{2}h_0)$, we compute $A(\frac{1}{4}h_0)$ and then we can find a linear combination of these three values which differs from A_0 by an $O(h_0^3)$ term. Moreover, we don't need to consider only the sequence: $h_0, \frac{1}{2}h_0,\ \frac{1}{4}h_0, \ldots,$ but a general sequence $h_0, h_1, h_2, \ldots$ of values of h, where:

$$h_0 > h_1 > h_2 > \ldots h_s > 0 \tag{3}$$

In general, we can find a linear combination with the property:

$$\sum_{s=0}^{S} c_{s,s}A(h_s) = A_0 + O(h_0^{s+1}), \quad h \to 0 \tag{4}$$

The forming of such linear combinations is essentially equivalent to polynomial interpolation at $h = 0$ of the data $(h_s, A(h_s))$, $s = 0, 1, 2, \ldots, S$. If the extrapolation is performed in an iterative manner (due originally to Aitken and Neville), it is avoid computing the coefficients $c_{s,S}$ in (4). For each value of h_s compute $A(h_s)$ and denote the results by $a_s^{(0)}$. Let $I_{01}(h)$ be the unique polynomial of degree 1 in h which interpolates the points $(h_0, a_0^{(0)})$ and $(h_1, a_1^{(0)})$ in the $h - A(h)$ plane. This polynom may be represented in terms of a 2×2 determinant as follows:

$$I_{01}(h) = \frac{1}{h_1 - h_0}\begin{vmatrix} a_0^{(0)} & h_0 - h \\ a_1^{(0)} & h_1 - h \end{vmatrix} \tag{5}$$

So defined $I_{01}(h)$ is indeed a polynomial of degree 1 in h and: $I_{01}(h_0) = a_0^{(0)}$, $I_{01}(h_1) = a_1^{(0)}$.

Let us denote by $a_0^{(1)}$ the result of extrapolating to $h = 0$ using this polynom; so $I_{01}(0) = a_0^{(1)}$. From (1) it follows that $a_0^{(1)} = A_0 + O(h_0^2)$, in the case when $h_1 = \frac{1}{2}h_0$; $a_0^{(1)}$ coincides with the left side of (2). We obtain a value $a_1^{(1)} = A_0 + O(h_1^2)$ by a similarly extrapolating to zero using the linear interpolant of the data $(h_1, a_1^{(0)})$ and $(h_2, a_2(0))$, where

$$\begin{aligned} a_1^{(1)} &= I_{12}(0), \\ I_{12}(h) &= \frac{1}{h_2 - h_1}\begin{vmatrix} a_1^{(0)} & h_1 - h \\ a_2^{(0)} & h_2 - h \end{vmatrix}. \end{aligned}$$

Now we note by $I_{012}(h)$ the unique polynomial of degree 2 which interpolates the points: $(h_0, a_0^{(0)})$, $(h_1, a_1^{(0)})$, $(h_2, a_2^{(0)})$ in the $h - A(h)$ plane. Then we may write:

$$I_{012}(h) = \frac{1}{h_2 - h_0} \begin{vmatrix} I_{01}(h) & h_0 - h \\ I_{12}(h) & h_2 - h \end{vmatrix}$$

since:

(i) $I_{012}(h)$ is a polynomial of degree 2 in h;
(ii) $I_{012}(h_0) = I_{01}(h_0) = a_0^{(0)}$; $I_{012}(h_2) = I_{12}(h_2) = a_2^{(0)}$;
(iii) $I_{012}(h_1) = [(h_2 - h_1) \cdot a_1^{(0)} - (h_0 - h_1)a_1^{(0)}]/(h_2 - h_0) = a_1^{(0)}$.

Extrapolating to zero by this polynomial define us: $a_0^{(2)} = I_{012}(0)$ and we find that $a_0^{(2)} = A_0 + O(h_0^3)$. Note that $a_0^{(2)}$ is a linear combination of $a_0^{(0)}$, $a_1^{(0)}$ and $a_2^{(0)}$, so that we have found the required linear combination (4) in the case $S = 2$. This process can be continued to give higher-order approximations to A_0 and we may be summarized by the following table:

$$\begin{array}{lllll} h_0 & a_0^{(0)} & & & \\ h_1 & a_1^{(0)} & a_0^{(1)} & & \\ h_2 & a_2^{(0)} & a_1^{(1)} & a_0^{(2)} & \\ h_3 & a_3^{(0)} & a_2^{(1)} & a_1^{(2)} & a_0^{(3)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{array} \tag{6}$$

where

$$\begin{aligned} a_s^{(0)} &= A(h_s), & s = 0, 1, 2, \ldots \\ a_s^{(m)} &= \tfrac{1}{h_{m+s} - h_s} \begin{vmatrix} a_s^{(m-1)} & h_s \\ a_{s+1}^{(m-1)} & h_{m+s} \end{vmatrix}, & m = 1, 2, \ldots \end{aligned} \tag{7}$$

For computation is more suitable an equivalent form of (7):

$$a_s^{(0)} = A(h_s); \quad a_s^{(m)} = a_{s+1}^{(m-1)} + \frac{a_{s+1}^{(m-1)} - a_s^{(m-1)}}{h_s/h_{m+s} - 1}, m = 1, 2, \ldots; \quad s = 0, 1, 2, \ldots \tag{8}$$

We show that: $a_s^{(m)} = A_0 + O(h_s^{m+1})$.

The benefits of repeated extrapolation are greatly enhanced if it happens that the asymptotic expansion for $A(h)$ contains only even powers of h - this will be the case for some important applications to ordinary differential equations.

For example, corresponding to (2), we have:

$$\frac{4}{3}A\left(\frac{1}{2}h_0\right) - \frac{1}{3}A(h_0) = A_0 + O(h_0^4).$$

If the asymptotic expansion for $A(h)$ has the form:

$$A(h) \approx A_0 + A_2h^2 + A_4h^4 + \dots \tag{9}$$

then the process of repeated polynomial extrapolation is described by (6), where now in place of (8):

$$a_s^{(0)} = A(h_s); \quad a_s^{(m)} = a_{s+1}^{(m-1)} + \frac{a_{s+1}^{(m-1)} - a_s^{(m-1)}}{(h_s/h_{m+s})^2 - 1}, m = 1, 2, \dots; \quad s = 0, 1, 2, \dots \tag{10}$$

We then have $a_s^{(m)} = A_0 + O(h_s^{2m+2})$.

The algorithm defined by (6) and (10), where (9) is assumed, is analysed by Gragg who shows that if $A(h)$ is continuous from the right at $h = 0$, then a necessary and sufficient condition for the convergence of $\{a_0^{(n)}\}$ to A_0 as $n \to \infty$ is that $\sup\limits_{n\geq 0}(h_{n+1}/h_n) < 1$. Gragg show that each column of (6) then converges to A_0 faster than the one to its left, and that if: $\inf\limits_{n\geq 0}(h_{n+1}/h_n) > 0$. The principal diagonal $a_0^{(0)}$, $a_0^{(1)}, a_0^{(2)}, \dots$ converges to A_0 faster than any column. Under mild conditions on $A(h)$, $\{a_0^{(n)}\}$ converges to A_0 superlinearly, in sens that: $|a_0^{(n)} - A_0| \leq K_n$ and $\lim\limits_{n\to\infty}(K_{n+1}/K_n) = 0$.

2 Application to initial value problems in ordinary differential equations

For a given discrete numerical method (linear multistep, Runge-Kutta) let denote $y(x;h)$ the approximation at x, given by the numerical method with steplength h, to the theoretical solution $y(x)$ of the initial value problem: $y' = f(x, y)$, $y(x_0) = y_0$.

We intend to use polynomial extrapolation to furnish approximations to $y(x)$ at the basic points $x_0 + jH$, $j = 0, 1, \dots$, where H is the basic steplength.

First we choose a steplength $h_0 = H/N_0$, where N_0 is a positive integer and apply the numerical method N_0 times starting from $x = x_0$ to obtain an approximation $y(x_0+H;h_0)$ to the theoretical solution $y(x_0+H)$. A second steplength $h_1 = H/N_1$, where N_1 is a positive integer greater than N_0 and the method applied N_1 times, again starting from $x = x_0$, to yield the approximation $y(x_0 + H; h_1)$.

Proceeding in this fashion for the sequence of steplengths $\{h_s\}$, where $h_s = H/N_s$, $\{N_s \mid s = 0, 1, \dots, S\}$ being an increasing sequence of positive integers, we obtain the sequence of approximations $\{y(x_0 + H; h_s) \mid s = 0, 1, \dots, S\}$ to $y(x_0 + H)$. In practice S is in the rand 4 to 7.

If, for the given numerical method, exists an asymptotic expansion of the form:

$$y(x;h) \sim y(x) + A_1 h + A_2 h^2 + A_3 h^3 + \dots \tag{11}$$

then we can set $a_s^{(0)} = y(x_0 + H; h_s)$ in (6) and apply the process of repeated polynomial extrapolation using (8). Equation (10) replaces (8) in the case when the numerical method possesses an asymptotic expansion of the form:

$$y(x;h) \sim y(x) + A_2 h^2 + A_4 h^4 + A_6 h^6 + \dots \tag{12}$$

Then we take the last entry in the main diagonal of the table (6) as our final approximation to $y(x_0 + H)$ and denote it by $y^*(x_0 + H; H)$.

To obtain a numerical solution at the next basic point x_0+2H, we apply the whole of the above procedure to the new initial value problem: $y' = f(x, y)$, $y(x_0 + H) = y^*(x_0 + H; H)$.

The motivation for extrapolation methods depends heavily on the possibility of choosing H to be large. Nevertheless, these results can be fairly described as generally encouraging, particularly since the numerical methods behind the algorithm, the mid-point rule, has no interval of absolute stability.

3 Weak stability

The success of extrapolation methods thus hinges on the existence of numerical methods which asymptotic expansions of the form (11) or, preferably (12). The existence of such expansions had frequently and rigorously investigated by Gragg.

Gragg's method (the modified mid-point method) is thus defined as follows:

$$\begin{aligned} &h_s = H/N_s, && N_s\text{-even}, \\ &y_0 = y(x_0), && \\ &y_1 = y_0 + h_s f(x_0, y_0) && \end{aligned} \tag{13}$$

$$\begin{aligned} &y_{m+2} - y_m = 2h_s \cdot f(x_{m+1}, y_{m+1}), \quad m = 0, 1, 2, \dots, N_s - 1, \\ &y(x_0 + H; h_s) = \frac{1}{4} y_{N_s+1} + \frac{1}{2} y_{N_s} + \frac{1}{4} y_{N_s-1}. \end{aligned}$$

If (13) is repeated for an increasing sequence N_s, $s = 0, 1, \dots, S$, of even integers, polynomial extrapolation using (6) and (10) can be applied. Two popular choices for the sequence $\{N_s\}$ are $\{2, 4, 6, 8, 12, 16, 24, \dots\}$ and $\{2, 4, 8, 16, 32, 64, \dots\}$.

We investigate the existence of asymptotic expansions of the form (12) for general implicit one-step methods of the form:

$$y_{n+1} - y_n = h \cdot \phi(x_n, x_{n+1}, y_n, y_{n+1}, h) \tag{14}$$

and we show that if the function ϕ satisfies the symmetry requirement:

$$\phi(s, t, \eta, \xi, h) = \phi(t, s, \xi, \eta, -h) \tag{15}$$

the (14) possesses an asymptotic expansion of the form (12). The trapezoidal rule satisfies this requirement, as does the implicit mid-point method,

$$y_{n+1} - y_n = h \cdot f\left(\frac{x_n + x_{n+1}}{2}, \frac{y_n + y_{n+1}}{2}\right).$$

This method suffers the disadvantage that it must be solved exactly for y_{n+1} at each step, if the asymptotic expansion is to remain valid. For practical purposes (13) still remains easily the most appropriate numerical method on which to base an extrapolation algorithm.

The mid-point method has no interval of absolute stability, but the algorithm has a non-vanishing interval of absolute stability. We cannot deduce from these intervals any useful information on the weak stability properties of the overall method consisting of (13), followed by polynomial extrapolation, since the latter forms linear combinations of the $y(x_0 + H; h_s)$ $s = 0, 1, 2, \ldots, S$ the coefficients in the combinations sometimes being negative.

Stetter adopts a new approach by computing the perturbation at $x_0 + H$ which results on introducing unit perturbations in each step of (13) applied to the linear differential equation $y' = \lambda y$, and then forming the linear combinations which correspond to the extrapolation process.

Now, with polynomial extrapolation for the sequence $N_s = 2, 4, 6, 8, 12$ (applying method (13)) we solve the initial value problem: $y' = -y$, $y(0) = 1$, for one basic step of length $H = 1,0$.

The table (6) can be constructed using (10). Each of the entries $a_s^{(m)}$ is an approximation to the theoretical solution $y(1) = e^{-1}$; the entries in brackets denote the global error: $y(1) - a_s^{(m)}$, multiplied by 10^5.

The results are clearly consistent on the rates of convergence of the columns and the principal diagonal. Note that the error in the final extrapolated value is roughly 40.000 times smaller than the error in the most accurate of the original computed solutions $y(1; h_s)$.

$h_s = H / N_s$	$a_s^{(m)}$				
$h_0 = \frac{1}{2}$	0,375000 (-712,056)				
$h_1 = \frac{1}{4}$	0,371093 (-321,431)	0,369791 (-191,223)			
$h_2 = \frac{1}{6}$	0,369455 (-157,644)	0,368145 (-26,614)	0,367939 (-6,038)		
$h_3 = \frac{1}{8}$	0,368796 (-91,739)	0,367949 (-7,004)	0,367884 (-0,467)	0,367880 (-0,096)	
$h_4 = \frac{1}{12}$	0,368297 (-41,768)	0,367897 (-1,791)	0,36787998 (-0,054)	0,36787946 (-0,002)	0,36787943 (0,001)

The first application of (13) clearly costs $N_0 + 1$ evaluations of the function $f(x, y)$. Each application of (13) for given N_s will cost N_s evaluations (since the evaluation of $f(x_0, y_0)$ need not be repeated). Thus, to compute $y(1; h_s)$ for the first 3, 4 and 5 members of sequence $\{\frac{1}{2}, \frac{1}{4}, \frac{1}{6}, \frac{1}{8}, \frac{1}{12}\}$ of steplength, costs respectively 13, 21 and 33 evaluations of f. Since the fourth-order Runge-Kutta method costs 4 evaluations per step, we shall compute solutions by it, using steplengths $\frac{1}{3}$, $\frac{1}{5}$ and $\frac{1}{8}$, which will costs 12, 20 and 32 evaluations respectivelly. We compared the errors in the two processes:

Polynomial extrapolation		Runge-Kutta	
Evaluations	Error	Evaluations	Error
13	$-6{,}038 \cdot 10^{-5}$	12	$-5{,}002 \cdot 10^{-5}$
21	$-0{,}096 \cdot 10^{-5}$	20	$-0{,}580 \cdot 10^{-5}$
33	$0{,}001 \cdot 10^{-5}$	32	$-0{,}083 \cdot 10^{-5}$

The superiority of the extrapolation method asserts itself only when $S \geq 2$. In practice S is typically in the rang 4 to 7.

References

[1] L.C. Piccinini, G. Stampacchia, G. Vidossich, Ordinary Differential Equations in $\mathbb{R}^n$. Problems and Methods. Springer-Verlag (New York, Berlin, Heidelberg, Tokyo).

[2] J.D. Lambert, Computational Methods in Ordinary Differential Equations, John Wiley&Sons (London, New York, Sydney, Toronto).

[3] Corneliu Berbente, Sorin Mitran, Silviu Zancu, Metode numerice, Editura Tehnică Bucureşti.