2011 International Symposium on Nonlinear Theory and its Applications
NOLTA2011, Kobe, Japan, September 4-7, 2011

On Discrete Approximation of Ordinary Differential Equations

Takatomi Miyata† and Masachika Miyata

† Kanazawa Institute of Technology
7-1 Ohgigaoka Nonoichi, Ishikawa, 921-8501 Japan
Email: takatomi-miyata@neptune.kanazawa-it.ac.jp, pulsar@goo.jp

Abstract– This paper shows typical examples to recall explicit Runge-Kutta methods are improper for long time integration, and novel expression of popular methods to explain why the implicit midpoint method is preferable from the view point of the correspondence between an analog model and a discrete model.

1. Introduction

The explicit/implicit Runge-Kutta methods are applied frequently to solve ordinary differential equations (ODEs) [1]-[3]. However it is well-known that confusing results are sometimes obtained by explicit methods, and that a simple implicit Runge-Kutta method which is equivalent to the midpoint method is always reasonable to solve ODEs, although it may be inconvertible exactly to an equivalent explicit method.

The purpose of this paper is to try *plain explanation* for students why explicit methods are improper for long time integration, and why the midpoint method is preferable, using iterative equations of discrete approximation.

2. Primitive Approximation

2.1. Preliminaries

In order to discuss inherent problems concerning with discrete approximation of ordinary differential equations, we use the following simple equation

$$\begin{pmatrix} x_1'(t) \\ x_2'(t) \end{pmatrix} = \begin{pmatrix} f_1(t,\ x_1(t),\ x_2(t)) \\ f_2(t,\ x_1(t),\ x_2(t)) \end{pmatrix} \tag{1}$$

or

$$\boldsymbol{x}'(t) = \boldsymbol{f}(t,\ \boldsymbol{x}(t)) \tag{2}$$

assuming that $\boldsymbol{f}(t, \boldsymbol{v})$ is differentiable with respect to $\boldsymbol{v}$ and that $\boldsymbol{x}(t)$ has Taylor expansion like

$$\boldsymbol{x}_{n+1} = \sum_{k=0}^{m} \frac{1}{k!} \boldsymbol{x}_n^{(k)} h^k + \boldsymbol{o}(h^m), \quad \lim_{h\to 0} \frac{\boldsymbol{o}(h^m)}{h^m} = \boldsymbol{0} \tag{3}$$

where

$$\boldsymbol{x}_n^{(k)} = \frac{d^k}{dt^k} \boldsymbol{x}(nh), \quad \boldsymbol{x}_n = \boldsymbol{x}_n^{(0)},\ \boldsymbol{x}_n' = \boldsymbol{x}_n^{(1)}. \tag{4}$$

The reason why we use little-o notation instead of big-O notation is that $\boldsymbol{x}_n + \boldsymbol{o}(h^2)$ may be $\boldsymbol{x}_n + \boldsymbol{o}(h^3)$.

The following is well-known examples to emphasize the problems in long time integration.

Example 1: Let $\boldsymbol{x}(t)$ be the solution of

$$\boldsymbol{x}'(t) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \boldsymbol{x}(t), \quad \boldsymbol{x}_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Then $\boldsymbol{x}(t) = (\cos t,\ \sin t)^{\mathrm{T}}$.

(a) The second order explicit Runge-Kutta method

$$\boldsymbol{x}_{n+1} = \boldsymbol{x}_n + h\boldsymbol{f}\left(nh + \frac{h}{2},\ \boldsymbol{x}_n + \frac{h}{2}\boldsymbol{f}(nh,\ \boldsymbol{x}_n)\right)$$

leads to

$$\boldsymbol{x}_{n+1} = \begin{pmatrix} 1 - h^2/2 & -h \\ h & 1 - h^2/2 \end{pmatrix} \boldsymbol{x}_n$$

$$\|\boldsymbol{x}_{n+1}\| = \left(1 + \frac{h^4}{4}\right) \|\boldsymbol{x}_n\|$$

Hence $\|\boldsymbol{x}_n\| \to \infty$ when $n \to \infty$.

(b) A one stage implicit Runge-Kutta method

$$\boldsymbol{x}_{n+1} = \boldsymbol{x}_n + h\boldsymbol{k}_1, \quad \boldsymbol{k}_1 = \boldsymbol{f}\left(nh + \frac{h}{2},\ \boldsymbol{x}_n + \frac{h}{2}\boldsymbol{k}_1\right)$$

which is equivalent to the midpoint method leads to

$$\boldsymbol{x}_{n+1} = \begin{pmatrix} 1 & h/2 \\ -h/2 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 & -h/2 \\ h/2 & 1 \end{pmatrix} \boldsymbol{x}_n$$

$$\|\boldsymbol{x}_{n+1}\| = \|\boldsymbol{x}_n\|$$

Hence $\|\boldsymbol{x}_n\| = 1$ for each n.

Example 2: Let $x'(t) = -ax(t)$ $(a > 0)$. Approximation such that

$$\frac{x(nh+h) - x(nh-h)}{2h} = -ax(nh)$$

seems to be better than

$$\frac{x(nh+h) - x(nh)}{h} = -ax(nh).$$

However the z-transform of $x(nh)$ using the former approximation is expressed as

$$X(z) = \frac{b}{1 - \left(\sqrt{1 + a^2h^2} - ah\right) z^{-1}} + \frac{c}{1 + \left(\sqrt{1 + a^2h^2} + ah\right) z^{-1}}$$

where b and c are the constants determined by $x(0)$ and $x(h)$. If $c = 0$, then $x(h) = (\sqrt{1 + a^2h^2} - ah)\, x(0)$.

If $x(h) = x(0) + x'(0)h$, then $\lim_{n\to\infty} |x(nh)| = \infty$, because $c \neq 0$ although $|c| \ll |b|$.

2.2. Midpoint Methods

The forward difference $\boldsymbol{x}_{n+1} - \boldsymbol{x}_n$ is an approximation of $\boldsymbol{x}'(nh + h/2)\,h$, then we consider the approximation of the differential equation at the midpoint such that

$$\boldsymbol{x}'(\tau_n) = \boldsymbol{f}(\tau_n,\ \boldsymbol{x}(\tau_n)) \tag{5}$$

$$\tau_n = nh + \frac{h}{2} \tag{6}$$

using the following operators D_n and A_n defined by

$$\mathrm{D}_n \boldsymbol{x}(t) = \frac{\boldsymbol{x}_{n+1} - \boldsymbol{x}_n}{h} \tag{7}$$

$$\mathrm{A}_n \boldsymbol{x}(t) = \frac{\boldsymbol{x}_{n+1} + \boldsymbol{x}_n}{2} \tag{8}$$

A primitive approximation of (5) is

$$\mathrm{D}_n \boldsymbol{x}(t) = \boldsymbol{f}(\tau_n,\ \mathrm{A}_n \boldsymbol{x}(t)) \tag{9}$$

since $\boldsymbol{x}'(\tau_n) = \mathrm{D}_n \boldsymbol{x}(t) + \boldsymbol{o}(h)$ and $\boldsymbol{x}(\tau_n) = \mathrm{A}_n \boldsymbol{x}(t) + \boldsymbol{o}(h)$. These operators are also used for scalar functions of t. For example, $\tau_n = \mathrm{A}_n t$.

Hereafter we distinguish so-called midpoint methods as follows:

MP1: $\mathrm{D}_n \boldsymbol{x}(t) = \boldsymbol{f}(\mathrm{A}_n t,\ \mathrm{A}_n \boldsymbol{x}(t))$

MP2: $\mathrm{D}_n \boldsymbol{x}(t) = \mathrm{A}_n \boldsymbol{f}(t,\ \boldsymbol{x}(t))$

MP3: $\dfrac{\boldsymbol{x}_{n+1} - \boldsymbol{x}_{n-1}}{2h} = \boldsymbol{f}(nh,\ \boldsymbol{x}_n)$

MP1 is an implicit Runge-Kutta method as in Example 1. MP2 is the most popular method and called trapezoidal method, or implicit modified Euler method. MP3 is called central difference method. The major difference between MP1 and MP2 is clarified by the following example.

Example 4: Consider the differential equation

$$\boldsymbol{x}'(t) = \begin{pmatrix} 0 & -2t \\ 2t & 0 \end{pmatrix} \boldsymbol{x}(t), \quad \boldsymbol{x}(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

of $\boldsymbol{x}(t) = (\cos t^2,\ \sin t^2)^{\mathrm{T}}$.

(a) The iteration by MP1 is

$$\boldsymbol{x}_{n+1} = \begin{pmatrix} 1 & h\tau_n \\ -h\tau_n & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 & -h\tau_n \\ h\tau_n & 1 \end{pmatrix} \boldsymbol{x}_n$$

$$\|\boldsymbol{x}_{n+1}\| = \|\boldsymbol{x}_n\|$$

(b) The iteration by MP2 is

$$\boldsymbol{x}_{n+1} = \begin{pmatrix} 1 & h(n+1) \\ -h(n+1) & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 & hn \\ hn & 1 \end{pmatrix} \boldsymbol{x}_n$$

$$\|\boldsymbol{x}_{n+1}\| = \frac{1 + h^2 n^2}{1 + h^2 (n+1)^2} \|\boldsymbol{x}_n\|$$

2.3. Comparison with explicit methods

The reason why explicit Euler method is worse than MP1 in Example 1 can be explained as follows. For simplicity, consider $y(t) = x_1(t) + i\,x_2(t)$ instead of $\boldsymbol{x}(t)$.

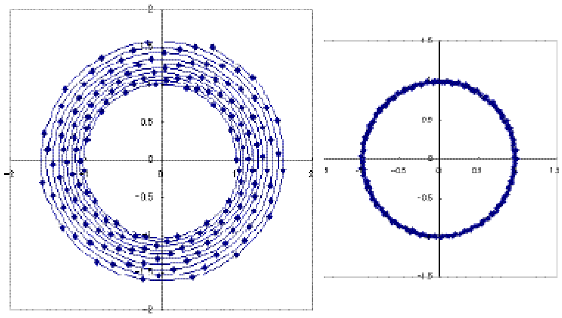
The exact iteration of $y'(t) = iy(t)$ is $y_{n+1} = e^{ih} y_n$ and its approximation by an explicit Runge-Kutta is

$$y_{n+1} = \sum_{k=0}^{m} \frac{1}{k!} (ih)^k y_n . \tag{10}$$

since $f(t, v) = iv$. Therefore, if m is even, then $y_n \to \infty$ $(n \to \infty)$ as shown in Figure 1(a), and if m is odd, then $y_n \to 0$ $(n \to \infty)$. Figure 1(b) shows an improvement of the solution using random switching of $m = 2$ and $m = 3$ such that

$$y_{n+1} = \left(1 + ih - \frac{h^2}{2} - i\frac{h^3}{6}\delta_n\right) y_n \tag{11}$$

where δ_n $(0 \le \delta_n \le 1)$ is a uniformly distributed random number



(a) y_n without dither (b) y_n with dither

FIGURE 1

COMPARISON OF TRAJECTORIES FOR $y(t) = e^{it}$

On the other hand, good solutions for linear ODEs are obtained by MP1 with sufficiently small h , the reason of which is explained later, referring the bilinear transform between s and z .

When $\boldsymbol{f}(t,\ \boldsymbol{v})$ is nonlinear, it is difficult to discuss approximation error in general. If any error bound of a method is unknown, a simple way to estimate it is to check the convergence of solutions for several h .

3. Linear approximation of MP1

Suppose that $\boldsymbol{f}(t,\ \boldsymbol{v})$ is nonlinear. Since MP1 is a low order approximation, a linear approximation of $\boldsymbol{f}(t,\ \boldsymbol{v})$ like

$$\begin{aligned} & f_i(t,\ v_1 + \Delta v_1,\ v_2 + \Delta v_2) \\ & \approx f_i(t,\ v_1,\ v_2) + \sum_{k=1}^{2} \Delta v_k \frac{\partial f_i}{\partial v_k}(t,\ v_1,\ v_2) \end{aligned} \tag{12}$$

does not degrade solutions so much as higher order approximation.

Example 3: Let $x(t) = \sqrt[3]{1+t^3}$. In this case,

$$x'(t) = f(t,\ x(t)), \quad f(t,\ v) = \frac{t^2}{v^2}$$

and $x'(\mathrm{A}_n t)$ is expressed as

$$x'(\mathrm{A}_n t) = f\left(\mathrm{A}_n t,\ x_n + \frac{h}{2}\mathrm{D}_n x(t) + o(h)\right)$$

$$f(t,\ v+\delta) = \frac{t^2}{v^2} - \frac{2t^2}{v^3}\delta + o(h).$$

Hence x_{n+1} of MP1 can be calculated as follows.

$$\frac{x_{n+1} - x_n}{h} = \frac{\tau_n^2}{x_n^2} - \frac{2\tau_n^2}{x_n^3} \cdot \frac{x_{n+1} - x_n}{2} \quad (\tau_n = \mathrm{A}_n t)$$

$$\left(1 + h\frac{\tau_n^2}{x_n^3}\right)x_{n+1} = \left(1 + 2h\frac{\tau_n^2}{x_n^3}\right)x_n .$$

The obtained solution for sufficiently small n is

$$x_{n+1} - x_n = h\frac{\tau_n^2}{x_n^2} + o(h^3) = h^3\left(n^2 + n + \frac{1}{4}\right) + o(h^3)$$

Since $n^2 + n + 1/4 \neq n^2 + n + 1/3$, this approximation error is $o(h^2)$ strictly, and rather $o(h^3)$ practically.

4. z-Transformed Linear Equations

When $\boldsymbol{f}(t,\ \boldsymbol{v})$ is linear with respect to t and $\boldsymbol{v}$, z-transform is available. For example, a simple MP1

$$\mathrm{D}_n x(t) = \mathrm{A}_n v(t) \tag{13}$$

is solved as follows.

$$\frac{z-1}{h}\sum_{n=0}^{\infty} x(nh)\, z^{-n} = \frac{z+1}{2}\sum_{n=0}^{\infty} v(nh)\, z^{-n} \tag{14}$$

$$D(z)X(z) - \frac{x(0)}{h} = A(z)V(z) - \frac{v(0)}{2} \tag{15}$$

$$X(z) = \frac{h}{2} \cdot \frac{1+z^{-1}}{1-z^{-1}} V(z) + \frac{x(0) - v(0)h/2}{1-z^{-1}} \tag{16}$$

where

$$\begin{aligned} X(z) &= \sum_{n=0}^{\infty} x(nh) z^{-n}, \quad V(z) = \sum_{n=0}^{\infty} v(nh) z^{-n} \\ D(z) &= \frac{1-z^{-1}}{h}, \quad A(z) = \frac{1+z^{-1}}{2} \end{aligned} \tag{17}$$

Example 4: Consider $\boldsymbol{x}(t)$ such that

$$x_1''(t) + a x_1'(t) + b x_1(t) = v(t) .$$

The MP1 equations are expressed as

$$\mathrm{D}_n x_1(t) = \mathrm{A}_n x_2(t)$$

$$\mathrm{D}_n x_2(t) = \mathrm{A}_n \left\{v(t) - a x_2(t) - b x_1(t)\right\}$$

which are z-transformed as

$$D(z)X_1(z) = A(z)X_2(z) + \alpha$$

$$D(z)X_2(z) = A(z)\left\{V(z) - aX_2(z) - bX_1(z)\right\} + \beta$$

$$\alpha = \frac{x_1(0)}{h} - \frac{x_2(0)}{2}$$

$$\beta = \frac{x_1(0)}{h} - \frac{v(0) - a x_2(0) - b x_1(0)}{2} .$$

Then the solution is obtained from

$$X_1(z) = \frac{A(z)^2 V(z) + A\beta + D\alpha + aA\alpha}{D(z)^2 + aD(z)A(z) + bA(z)^2} .$$

Let $V(z) = 1, \quad \alpha = \beta = 0$ and $h = 2$, and assume $x_1(t)$ is the impulse response of a system. Remark that $X_1(z)$ is equal to the bilinear transform of the analog transfer function [4], namely

$$X_1(z) = \frac{1}{s^2 + as + b} \quad \left(s = \frac{z-1}{z+1}\right)$$

This property holds in general, because MP1 like (13) corresponds to a bilinear transform $\dfrac{sh}{2} = \dfrac{z-1}{z+1}$.

Example 5: $X(z)$ of Example 2 is the solution of

$$A(z)D(z)X(z) = -az^{-1}X(z) + \alpha + \beta z^{-1} .$$

A correct z-transformed equation of MP3 is

$$A(z)D(z)X(z) = -aA(z)^2 X(z) + \frac{2+ah}{2h} A(z)x(0)$$

where the pole of $X(z)$ is only at $z = (2-ah)/(2+ah)$.

5. Conclusion

It is easy to solve linear ODEs analytically. However such examples suggest general correspondence between an analog model and its digital model. The midpoint method seems to be better than higher order explicit Runge-Kutta methods for long time integration, because it has reasonable base for linear ODEs related with the bilinear transform between s and z which is popular in digital signal processing.

Recent researches on higher order implicit methods as described in [3] are not discussed in this paper because it is beyond plain explanation.

References

[1] E. Hairer, S. P. Nørsett, and G. Wanner, *Solving ordinary differential equations I*, 2nd revised. ed., 1993, Springer-Verlag, New York, ISBN 0-387-56670-8.

[2] J. C. Butcher, "Implicit Runge-Kutta Processes", http://www.ams.org/journals/mcom/1964-18-085/S0025-5718-1964-0159424-9/S0025-5718-1964-0159424-9.pdf

[3] D. I. Ketcheson, C. B. Macdonald, and S. Gottlieb, "Optimal implicit strong stability preserving Runge-Kutta methods", http://www.cfm.brown.edu/people/sg/implicitSSP.pdf
(29 references)

[4] A. V. Oppenheim and R. W. Schafer, *Digital Signal Processing*, Prentice Hall, 1975, ISBN 978-0132146357.