Simulation Methods

Numerical Methods for Ordinary Differential Equations

Joan Carles Tatjer

Departament de Matemàtiques i Informàtica

Universitat de Barcelona

Existence and uniqueness of Solution of the Cauchy Problem I

Consider the one-dimensional differential equation with initial condtion (Cauchy problem):

$$y' = f(x, y), \quad y(x_0) = y_0.$$

To find $y(X)$ we consider a subdivision

$$x_0, x_1, \ldots, x_{n-1}, x_n = X$$

and repalce in each subinterval the solution by the first term of its Taylor series:

$$\begin{aligned}
y_1 &= y_0 + (x_1 - x_0) f(x_0, y_0), \\
y_2 &= y_1 + (x_2 - x_1) f(x_1, y_1), \\
&\ldots \\
y_n &= y_{n-1} + (x_n - x_{n-1}) f(x_{n-1}, y_{n-1}).
\end{aligned}$$

Existence and uniqueness of Solution of the Cauchy Problem II

We define $h = (h_0, h_1, \ldots, h_{n-1})$, where $h_i = x_{i+1} - x_i$, and the **Euler polygon**:

$$y_h(x) = y_i + (x - x_i) f(x_i, y_i) \qquad \text{for} \quad x_i \leq x \leq x_{i+1}.$$

Existence and uniqueness of Solution of the Cauchy Problem III

Lemma

Assume that $|f|$ is bounded by A on

$$D = \{(x, y) \mid x_0 \leq x \leq X, \ |y - y_0| \leq b\}.$$

If $X - x_0 \leq b/A$ then the numerical solution (x_i, y_i) given above, remains in D for every subdivision and we have

$$|y_h(x) - y_0| \leq A|x - x_0|,$$

$$|y_h(x) - (y_0 + (x - x_0)f(x_0, y_0))| \leq \epsilon|x - x_0|$$

if $|f(x, y) - f(x_0, y_0)| \leq \epsilon$ on D.

Proof:

Existence and uniqueness of Solution of the Cauchy Problem IV

The lemma is obviously true if $x = x_0$. Therefore, we take $x \in \mathbb{R}$ such that $x_0 < x \leq X$, and a subdivision. Then, there exists $1 < j \leq n$ such that $x_{j-1} < x \leq x_j$. Then

$$\begin{aligned}
y_1 - y_0 &= (x_1 - x_0) f(x_0, y_0), \\
y_2 - y_1 &= (x_2 - x_1) f(x_1, y_1), \\
&\cdots \\
y_{j-1} - y_{j-2} &= (x_{j-1} - x_{j-2}) f(x_{j-2}, y_{j-2}), \\
y_h(x) - y_{j-1} &= (x - x_{j-1}) f(x_{j-1}, y_{j-1}).
\end{aligned}$$

Then, by adding up and using the triangle inequality, we have

$$|y_h(x) - y_0| \leq A(x - x_0),$$

Existence and uniqueness of Solution of the Cauchy Problem V

For the other inequality, if $x_0 \leq x \leq x_1$ it is trivially true. If $x > x_1$

$$
\begin{aligned}
y_1 - y_0 &= (x_1 - x_0) f(x_0, y_0), \\
y_2 - y_1 &= (x_2 - x_1)(f(x_1, y_1) - f(x_0, y_0)) + (x_2 - x_1) f(x_0, y_0), \\
&\dots \\
y_{j-1} - y_{j-2} &= (x_{j-1} - x_{j-2})(f(x_{j-2}, y_{j-2}) - f(x_0, y_0)) + \\
&\quad + (x_{j-1} - x_{j-2}) f(x_0, y_0), \\
y_h(x) - y_{j-1} &= (x - x_{j-1})(f(x_{j-1}, y_{j-1}) - f(x_0, y_0)) + \\
&\quad + (x - x_{j-1}) f(x_0, y_0).
\end{aligned}
$$

$$
y_h(x) - y_0 = (x - x_0) f(x_0, y_0) + \sum_{k=1}^{j-2} (x_{k+1} - x_k)(f(x_k, y_k) - f(x_0, y_0)) +
$$

$$
+ (x - x_{j-1})(f(x_{j-1}, y_{j-1}) - f(x_0, y_0)).
$$

Existence and uniqueness of Solution of the Cauchy Problem VI

Finally,

$$|y_h(x) - y_0 - (x - x_0)f(x_0, y_0)| \leq (x - x_1)\epsilon \leq (x - x_0)\epsilon.$$

Using the first formula, we see that the polygon remains in D. $\square$

Now, we want to obtain an estimate for the change of $y_h(x)$, when the initial value is changed:

Lemma

For a fixed subdivision h let $y_h(x)$ and $z_h(x)$ be the Euler polygons corresponding to the initial values y_0 and z_0, respectively. If

$$\left|\frac{\partial f}{\partial y}(x, y)\right| \leq L$$

in a convex region which contains $(x, y_h(x))$ and $(x, z_h(x))$ for all $x_0 \leq x \leq X$, then

$$|z_h(x) - y_h(x)| \leq e^{L(x-x_0)}|z_0 - y_0|.$$

Proof:

We have

$$
\begin{aligned}
y_1 - y_0 &= (x_1 - x_0) f(x_0, y_0), \\
z_1 - z_0 &= (x_1 - x_0) f(x_0, z_0),
\end{aligned}
$$

Substracting the second equation from the first, we get

$$z_1 - y_1 = z_0 - y_0 + (x_1 - x_0)(f(x_0, z_0) - f(x_0, y_0)).$$

Now, by the Mean Value Theorem:

$$|f(x, z) - f(x, y)| \leq L|x - y|,$$

and, therefore,

$$|z_1 - y_1| \leq (1 + (x_1 - x_0)L)|z_0 - y_0| \leq e^{L(x_1 - x_0)}|z_0 - y_0|.$$

If we do the same for $z_2 - y_2$, we obtain

$$|z_2 - y_2| \leq e^{L(x_2 - x_1)}|z_1 - y_1| \leq e^{L(x_2 - x_0)}|z_0 - y_0|.$$

Repeating the same argument, we obtain the result. □

If $|h| = \max_{i=0,\dots,n-1} h_i \to 0$, the Euler polygons converge to the solution of the Cauchy problem:

Theorem

Let $f(x,y)$ be continuous, and $|f|$ be bounded by A and satisfy the Lipschitz condition

$$|f(x,z) - f(x,y)| \leq L|z-y|$$

on

$$D = \{(x,y) \mid x_0 \leq x \leq X,\ |y-y_0| \leq b\}.$$

If $X - x_0 \leq b/A$, then we have:

1. *For $|h| \to 0$ the Euler polygons $y_h(x)$ converge uniformly to a continuous function $\varphi(x)$.*
2. *$\varphi(x)$ is continuously differentiable and solution of the Cauchy Problem (CP). on $x_0 \leq x \leq X$.*
3. *There exists no other solution of the CP on $x_0 \leq x \leq X$.*

Proof:

a) Take $\epsilon > 0$. Since f is uniformly continuous on the compact set D, $\exists\, \delta > 0$ s.t.

$$|u_1 - u_2| \leq \delta \text{ and } |v_1 - v_2| \leq A\delta \quad \Rightarrow \quad |f(u_1, v_1) - f(u_2, v_2)| \leq \epsilon.$$

Suppose that the subdivision satisfies

$$|x_{i+1} - x_i| \leq \delta, \text{ that is } |h| \leq \delta.$$

Consider a subdivision $h(1)$, which is obtained by adding new points only to the first subinterval. From Lemma 1 (applied to the first interval)

$$|y_{h(1)}(x_1) - y_h(x_1)| \leq \epsilon |x_1 - x_0|.$$

Since the subdivisions h and $h(1)$ are identical on $x_1 \leq x \leq X$, we can apply Lemma 2 to obtain

$$|y_{h(1)}(x) - y_h(x)| \leq e^{L(x - x_1)}(x_1 - x_0)\epsilon \text{ for } x_1 \leq x \leq X.$$

Let $h(2)$ be a subdivision obtained adding to $h(1)$ points in (x_1, x_2). Then

$$|y_{h(2)}(x_2) - y_{h(1)}(x_2)| \leq \epsilon |x_2 - x_1|$$

and

$$|y_{h(2)}(x) - y_{h(1)}(x)| \leq e^{L(x-x_2)}(x_2 - x_1)\epsilon \text{ for } x_2 \leq x \leq X.$$

Then

$$|y_{h(2)}(x) - y_h(x)| \leq e^{L(x-x_1)}(x_1 - x_0)\epsilon + e^{L(x-x_2)}(x_2 - x_1)\epsilon, \text{ for } x_2 \leq x \leq X.$$

If we denote by $\hat{h}$ the final refinement, we obtain for $x_i < x \leq x_{i+1}$

$$|y_{\hat{h}}(x) - y_h(x)| \leq \epsilon(e^{L(x-x_1)}(x_1 - x_0) + \cdots + e^{L(x-x_i)}(x_i - x_{i-1})) + \epsilon(x - x_i) \leq$$

$$\leq \epsilon \int_{x_0}^{x} e^{L(x-s)}\, ds = \frac{\epsilon}{L}(e^{L(x-x_0)} - 1),$$

where we add the term $\epsilon(x - x_i)$ in order to be the inequality true also in the case $i = 0$.

If we now have two subdivisions h and $\tilde{h}$ s.t. $|h| \leq \delta$ and $|\tilde{h}| \leq \delta$, let $\hat{h}$ be a subdivision which is a refinement of both subdivisions. Then

$$|y_h(x) - y_{\tilde{h}}(x)| \leq 2\frac{\epsilon}{L}(e^{L(x-x_0)} - 1).$$

This implies uniform convergence of y_h, when $|h| \to 0$, and therefore convergence to a continuous function $\varphi(x)$.

b) Let

$$\epsilon(\delta) = \sup\{|f(u_1, v_1) - f(u_2, v_2)| \,/\, |u_1 - u_2| \leq \delta, |v_1 - v_2| \leq A\delta, (u_i, v_i) \in D\}$$

be the modulus of continuity. If x belongs to the subdivision h then we obtain from Lemma 1, replacing (x_0, y_0) by $(x, y_h(x))$ and x by $x + \delta$,

$$|y_h(x + \delta) - y_h(x) - \delta f(x, y_h(x))| \leq \epsilon(\delta)\delta.$$

Taking the limit $|h| \to 0$ we get

$$|\varphi(x + \delta) - \varphi(x) - \delta f(x, \varphi(x))| \leq \epsilon(\delta)\delta.$$

Since $\epsilon(\delta) \to 0$ for $\delta \to 0$, this proves the differentiability of $\varphi(x)$ and $\varphi'(x) = f(x, \varphi(x))$.

c) Let $\psi(x)$ be a second solution and suppose that $|h| \leq \delta$. We then denote by $y_h^{(i)}(x)$ the Euler polygon to the initial value $(x_i, \psi(x_i))$ (it is defined for $x_i \leq x \leq X$). we have

$$\psi(x) = \psi(x_i) + \int_{x_i}^{x} f(s, \psi(s))\, ds$$

and

$$|\psi(x) - y_h^{(i)}(x)| = \left| \int_{x_i}^{x} f(s, \psi(s))\, ds - (x - x_i) f(x_i, \psi(x_i)) \right| =$$

$$= \left| \int_{x_i}^{x} (f(s, \psi(s)) - f(x_i, \psi(x_i))\, ds \right| \leq \epsilon |x - x_i| \text{ for } x_i \leq x \leq x_{i+1}.$$

In particular, $y_h^{(0)} = y_h$. Therefore, taking the limits $|h| \to 0$ and $\epsilon \to 0$, we obtain $|\psi(x) - \varphi(x)| \leq 0$, for $x \in [x_0, x_1]$. If we repeat the argument for all i, we see that $\psi(x) = \varphi(x)$, for all $x_0 \leq x \leq X$. □

Comment

In the proof of part a) of the theorem, we see that

$$|y_{\tilde{h}}(x) - y_h(x)| \leq \frac{\epsilon}{L}(e^{L(x-x_0)} - 1).$$

If we take the limit $|\tilde{h}| \to 0$, *we obtain the following* **error estimate**

$$|y(x) - y_h(x)| \leq \frac{\epsilon}{L}(e^{L(x-x_0)} - 1),$$

for the Euler polygon ($|h| \leq \delta$)*. Here* $y(x)$ *stands for the exact solution of the Cauchy problem.*

To end this introduction, we can give a general theorem of existence and uniqueness of the solution of the Cauchy problem $y' = f(x, y)$, $y(x_0) = x_0$:

Theorem

Let $f(x,y)$ be continuous, $|f|$ be bound by A and satisfy the Lipschitz condition $|f(x,z) - f(x,y)| \leq L|z-y|$ on

$$D = \{(x,y) \,|\, x_0 \leq x \leq X, \, |y - y_0| \leq b\}.$$

If $X - x_0 \leq b/A$, then we have

1. *For $|h| \to 0$ the Euler polygons $y_h(x)$ converge uniformly to a continuous function $\varphi(x)$.*
2. *$\varphi(x)$ is continuously differentiable and solution of the Cauchy problem on $x_0 \leq x \leq X$.*
3. *There exists no other solution on $x_0 \leq x \leq X$.*
4. *If we suppose, moreover, that $|\partial f/\partial y| \leq L$, $|\partial f/\partial x| \leq M$ then*

$$|y(x) - y_h(x)| \leq \frac{M + AL}{L}(e^{L(x-x_0)} - 1)|h|,$$

provided that $|h|$ is sufficiently small.

Proof:
It remains only to prove item d). For $|u_1 - u_2| \leq |h|$ and $|v_1 - v_2| \leq A|h|$ we obtain the estimate

$$|f(u_1, v_1) - f(u_2.v_2)| \leq (M + AL)|h|.$$

When we insert $\epsilon = (M + AL)|h|$ in the proof of the previous theorem, we obtain the desired result. $\square$

Comment

In the case we have a system of ordinary differential equations, that is $f = (f_1, \ldots, f_n)$, $y = (y_1, \ldots, y_n)$ *we obtain the same theorem, replacing absolute values by norms.*

Overview of single step methods of integration I

We want to solve the Cauchy Problem (CP):

$$\left.\begin{array}{rcl} \dot{x} & = & f(t,x) \\ x(t_0) & = & x_0 \end{array}\right\}, \qquad x, f \in \mathbb{R}^n, \quad t \in \mathbb{R}.$$

Given a sequence $t_0 < t_1 < \cdots$, we want to approximate the solution of the CP at these values. Let $x(t) = \phi(t; t_0, x_0)$ be its solution. We look for approximations of the table:

$$\begin{array}{c|c} t & x \\ \hline t_0 & x_0 = \phi(t_0, t_0, x_0) \\ t_1 & x_1 = \phi(t_1, t_0, x_0) \\ t_2 & x_2 = \phi(t_2, t_0, x_0) \\ \vdots & \vdots \end{array},$$

Overview of single step methods of integration II

that is

$$\begin{array}{c|c} t & x \\ \hline t_0 & x_0 \\ t_1 & \tilde{x}_1 \\ t_2 & \tilde{x}_2 \\ \vdots & \vdots \end{array}.$$

We call $h_n = t_{n+1} - t_n$ the n-th **step size**.

Overview of single step methods of integration III

Let $\varphi(h;t,x) = \phi(t+h;t,x)$ be the map which gives the solution with initial condition $x(t) = x$ after h time units. Then, we can write:

$$\left.\begin{array}{rcl} t_{n+1} & = & t_n + h_n, \\ x_{n+1} & = & \varphi(h_n; t_n, x_n) \end{array}\right\}$$

We replace φ (unknown) by and approximation $\tilde{\varphi}(h;t,x)$, s. t. $\tilde{\varphi}(0;t,x) = \varphi(0;t,x) = x$. Then $(t_0, \tilde{x}_0) = (t_0, x_0)$ and for $n \geq 0$:

$$\left.\begin{array}{rcl} t_{n+1} & = & t_n + h_n, \\ \tilde{x}_{n+1} & = & \tilde{\varphi}(h_n; t_n, \tilde{x}_n) \end{array}\right\}$$

Comment

The map $\tilde{\varphi}(h;t,x)$ completely determines the numerical method. As $\tilde{\varphi}$ only uses $(t_n, \tilde{x}_n)$ and h_n to compute $(t_{n+1}, \tilde{x}_{n+1})$,, we call this kind of numerical methods, **one (or single) step methods**.

Overview of single step methods of integration IV

Euler's Method

As

$$\dot{x}(t) = f(t, x(t)),$$

and

$$\dot{x}(t) \approx \frac{x(t+h) - x(t)}{h},$$

we have that

$$x(t+h) \approx x(t) + hf(t, x(t)).$$

Then, we define $\tilde{\varphi}(h; t, x) = x + hf(t, x)$. The corresponding method is

$$\left.\begin{array}{rcc} t_{n+1} & = & t_n + h_n, \\ \tilde{x}_{n+1} & = & \tilde{x}_n + h_n f(t_n, \tilde{x}_n) \end{array}\right\}$$

Overview of single step methods of integration V

Consistency

We say that a single-step method with map $\tilde{\varphi} = \tilde{\varphi}(h; t, x)$ is **consistent** if

$$\frac{\partial \tilde{\varphi}}{\partial h}(0; t, x) = f(t, x).$$

Comment

1. *As*

$$\frac{\partial \varphi(h; t, x)}{\partial h} = f(t + h, \varphi(h; t, x)),$$

we have

$$\frac{\partial \varphi}{\partial h}(0; t, x) = f(t, x).$$

2. *The Euler's method is consistent.*

Overview of single step methods of integration VI

Local truncation error

The **local truncation error** in the n-th step of a single-step method associated to the function $\tilde{\varphi}(h; t, x)$ is

$$\tilde{x}_n - \phi(t_n; t_{n-1}, \tilde{x}_{n-1}).$$

Global truncation error

The **global truncation error** is

$$\tilde{x}_n - \phi(t_n; t_0, x_0).$$

Local order of convergence

If $h = h_n$, for all n and $\tilde{x}_n - \phi(t_n; t_{n-1}, \tilde{x}_{n-1}) = O(h^m)$ we say that the **local order of convergence** is m.

Overview of single step methods of integration VII

Global order of convergence

If $h = h_n$, for all n and $\tilde{x}_n - \phi(t_n; t_0, \tilde{x}_0) = O(h^m)$ we say that the **global order of convergence** is n.

Comment

In general, one can prove that if the local order of convergence is $m + 1$ *then the corresponding global order is* m, *when* $h \to 0$. *We have the same behaviour when we compare the simple and composite methods in numerical integration.*

Orders of convergence of the Euler's method I

We write

$$\begin{aligned} t_n &= t_{n-1} + h_{n-1}, \\ \tilde{x}_n &= \tilde{\varphi}(h_{n-1}, t_{n-1}, \tilde{x}_{n-1}), \\ x_n &= \varphi(h_{n-1}, t_{n-1}, \tilde{x}_{n-1}). \end{aligned}$$

Therefore, if we write $h = h_{n-1}$, $x = \tilde{x}_{n-1}$, $t = t_{n-1}$:

$$\tilde{x}_n - x_n = \tilde{\varphi}(h; t, x) - \varphi(h; t, x).$$

Now we perform the Taylor expansion at $h = 0$, taking into account that

$$\tilde{\varphi}(h; t, x) = x + hf(t, x), \qquad \varphi(h; t, x) = \phi(t + h; t, x).$$

We have

$$\tilde{\varphi}(0; t, x) = \varphi(0; t, x) = x,$$

Orders of convergence of the Euler's method II

$$\frac{\partial \tilde{\varphi}}{\partial h}(0;t,x) = \frac{\partial \varphi}{\partial h}(0;t,x) = f(t,x) \quad \text{(consistency)},$$

and

$$\frac{\partial^2 \tilde{\varphi}}{\partial h^2}(0;t,x) = 0, \qquad \frac{\partial^2 \varphi}{\partial h^2}(0;t,x) = \frac{\partial f}{\partial t}(t,x) + \frac{\partial f}{\partial x}(t,x) f(t,x).$$

We obtain the latter derivative when we differentiate the expression

$$\frac{\partial \varphi}{\partial h}(h;t,x) = f(t+h, \varphi(h;t,x)),$$

with respect to h, and take $h = 0$. The two derivatives are different in general, which means that the local order of convergence is **2**. We have seen before that the global order of convergence is **1**.

Explicit and implicit one-step methods

The methods we have seen are **explicit methods**. Suppose that we have the following method:

$$\left.\begin{array}{rcl} t_{n+1} & = & t_n + h_n, \\ \tilde{x}_{n+1} & = & \hat{\varphi}(h_n; t_n, \tilde{x}_n, \tilde{x}_{n+1}) \end{array}\right\}$$

beginning with the initial condition $(t_0, \tilde{x}_0) = (t_0, x_0)$. This is an **implicit method** and can be transformed into an explicit one if we can isolate $\tilde{x}_{n+1}$ from the last equation, that is the function $\tilde{\varphi}$ satisfies:

$$\tilde{\varphi}(h; t, x) = \hat{\varphi}(h; t, x, \tilde{\varphi}(h; t, x)).$$

The Implicit Euler's Method I

We take

$$\dot{x}(t) \approx \frac{x(t) - x(t-h)}{h},$$

then, as $\dot{x}(t) = f(t, x(t))$, we have that

$$\frac{x(t) - x(t-h)}{h} \approx f(t, x(t)),$$

or

$$x(t+h) \approx x(t) + hf(t+h, x(t+h)).$$

Now, we can define the method generated by the function

$$\tilde{\varphi}(h; t, x) = x + hf(t+h, \tilde{\varphi}(h; t, x)).$$

Therefore, we have

$$\left.\begin{array}{rcc} t_{n+1} & = & t_n + h_n \\ \tilde{x}_{n+1} & = & \tilde{x}_n + h_n f(t_{n+1}, \tilde{x}_{n+1}) \end{array}\right\} \quad \textbf{(Implicit Euler's Method)}$$

The Implicit Euler's Method II

Implementation of the Implicit Euler's Method

To compute $\tilde{x}_{n+1}$ we have to solve a (in general) nonlinear equation. In each step:

- Obtain an approximation $\tilde{x}_{n+1}^{(0)}$ of $\tilde{x}_{n+1}$, using the (explicit) Euler's Method.
- Define the sequence $(\tilde{x}_{n+1}^{(i)})_i$, such that

$$\tilde{x}_{n+1}^{(i+1)} = \tilde{x}_n + h_n f(t_{n+1}, \tilde{x}_{n+1}^{(i)}), \quad i \geq 0,$$

  and take $\tilde{x}_{n+1} = \tilde{x}_{n+1}^{(j)}$, for some j small.

The Implicit Euler's Method III

Comment

Note that the function $F(x) = \tilde{x}_n + hf(t_{n+1}, x)$ *is a contraction if* $|h|$ *is small enough, which implies that if* $f : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ *is continuous and Lipschitz with respect to* x*, then* $\tilde{x}_{n+1}^{(i)} \to \tilde{x}_{n+1}$ *when* $i \to \infty$*. However, in the practical use of this method we will only consider a small number of iterations.*

Comment

This implemetation is called a **predictor-corrector method.**

Order of convergence of the Implicit Euler's Method I

We have

$$\tilde{\varphi}(h;t,x) = x + hf(t, \tilde{\varphi}(h;t,x)).$$

Then we write

$$\tilde{\varphi}(h;t,x) = c_0 + c_1 h + c_2 h^2 + \cdots ,$$

where $c_i = c_i(t,x)$.

- We insert this identity in the previous equation and take $h = 0$, we obtain that $c_0 = x$.
- Then we write

$$c_1 + c_2 h + O(h^2) = f(t, x + c_1 h + c_2 h^2 + O(h^2)) =$$

$$= f(t,x) + D_2 f(t,x) c_1 h + O(h^2).$$

  obtaining $c_1 = f(t,x)$.

Order of convergence of the Implicit Euler's Method II

- Finally,

$$f(t,x) + c_2 h + O(h^2) = f(t, x + f(t,x)h + c_2 h^2 + O(h^2)) =$$

$$f(t,x) + D_2 f(t,x) f(t,x) h + O(h^2),$$

which implies that $c_2 = D_2 f(t,x) f(t,x)$ and

$$\tilde{\varphi}(h; t, x) = x + f(t,x)h + \frac{\partial f}{\partial x}(t,x) f(t,x) h^2 + O(h^3).$$

- As

$$\varphi(h; t, x) = x + f(t,x)h + \frac{1}{2}\left(\frac{\partial f}{\partial t}(t,x) + \frac{\partial f}{\partial x}(t,x) f(t,x)\right) h^2 + O(h^3)$$

then

$$\tilde{\varphi}(h; t, x) - \varphi(h; t, x) = O(h^2).$$

Therefore, the local error is of order **2**, and the global error is of order **1**.