Universitatea Babeş-Bolyai, Facultatea de Matematică şi Informatică
Secţia: Informatică engleză, Curs: Dynamical Systems, Semestru: Primăvara 2018

Chapter 5. Numerical methods for differential equations[1]

We consider the IVP for a scalar first order differential equation

$$y' = f(x, y), \quad y(x_0) = y_0. \tag{1}$$

where $f : \mathbb{R}^2 \to \mathbb{R}$ is C^1 and $(x_0, y_0) \in \mathbb{R}^2$ is fixed. Here the unknown is the function y of variable x, and y' denotes its derivative with respect to x. We have the following result

The IVP (1) *has a unique solution denoted* φ*, defined at least on an interval* $[x_0, x^*]$ *for* $x^* > x_0$.

As we already know, not always it is possible to find the exact expression of the solution φ. Because of this, a theory on how to find *good* approximations of φ have been developed. The *numerical methods* are part of this theory. Their aim is to find approximations for the values of the solution on some given points in the interval $[x_0, x^*]$. More precisely, if we consider a partition of the interval $[x_0, x^*]$,

$$x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = x^*,$$

the purpose is to find some values denoted y_k as good approximations of $\varphi(x_k)$, for any $k = \overline{1, n}$. Then an approximate solution (i.e. function) can be found using interpolation methods (these type of methods are able to find a smooth function whose graph passes through the known points (x_k, y_k) , for any $k = \overline{1, n}$).

The approximate values y_k are usually computed using a recurrence formula. There are now many such formulas, many of them adapted to particular classes of equations or systems. We will present here only the basic ones: the Euler method (discovered by the Swiss mathematician Leonhard Euler around 1765) and the improved Euler's method.

For simplicity we work only with partitions of the interval where the points are at equal distance $h > 0$, i.e. they satisfy, for any $k = \overline{0, n-1}$,

$$x_{k+1} = x_k + h.$$

[1]©2018 Adriana Buică, *Linear Differential Systems*

One can deduce that $x_k = x_0 + k\,h$ for any $k = \overline{1,n}$. The number n is called *the number of steps* to reach the end of the interval. When given a step size $h > 0$ and an interval $[x_0, x^*]$, the number of steps to arrive as close to x^* as possible, is

$$n = \left[\frac{x^* - x_0}{h}\right]$$

where $[\cdot]$ denotes the entire part.

When given the number of steps $n \in \mathbb{N}^*$ and an interval $[x_0, x^*]$, the step size is

$$h = \frac{x^* - x_0}{n}.$$

The Euler's method formula for the IVP (1) *with constant step size h is*

$$y_{k+1} = y_k + h\,f(x_k, y_k), \quad k = \overline{0, n-1}.$$

The improved Euler's method formula for the IVP (1) *with constant step size h is*

$$y_{k+1} = y_k + \frac{h}{2}\,f(x_k, y_k) + \frac{h}{2}\,f(x_{k+1}, y_k + h\,f(x_k, y_k)), \quad k = \overline{0, n-1}.$$

Note that the starting point y_0 is the one that appears in the initial condition in (1). Hence $y_0 = \varphi(x_0)$ is an exact value. In fact, only in theory y_0 is an exact value, since practically, when y_0 has too many decimals (for example is an irrational number) a human or even a computer use only a truncation of it. When applying a numerical method, the errors are due to the formula itself and to the truncations made. Moreover, the errors accumulate at each step, thus, in general, *the errors are larger as the interval $[x_0, x^*]$ is larger.*

When the step size h is smaller, the partition of the interval $[x_0, x^*]$ is finer. In general *the errors are smaller as the step size h is smaller.*

Exercise 1. We consider the IVP $y' = y,\ y(0) = 1$ whose solution we know that it is $\varphi : \mathbb{R} \to \mathbb{R}$, $\varphi(x) = e^x$. Apply the Euler numerical method with a constant step size $h > 0$ on the interval $[0, x^*]$ where $x^* > 0$ is fixed. Prove that

$$y_k = (1+h)^k, \quad k = \overline{0,n} \quad \text{where } h = \frac{x^*}{n}.$$

Prove that $y_n \to \varphi(x^*) = e^{x^*}$ as $n \to \infty$. $\diamond$

Exercise 2. We consider the IVP $y' = 1 + xy^2,\ y(0) = 0$ whose unique solution is denoted by φ. Write the two numerical formulas with constant step size $h > 0$ for this IVP. Now take $h = 0.1$. Find the number of steps to reach $x^* = 1$. For each of the two formulas, compute approximate values for $\varphi(0.1)$, $\varphi(0.2)$ and $\varphi(0.3)$. $\diamond$

In the rest of the lecture we present two ideas on how the Euler's numerical formula can be derived.

The first idea uses the notion of *Taylor polynomial.* We know that the Taylor polynomial around a point a of some function φ is a good approximation of it at least in a small neighborhood of a. The highest the degree of the Taylor polynomial, the better the approximation. But we consider only the Taylor polynomial of degree 1, that is

$$\varphi(a) + (x - a)\varphi'(a).$$

With this we approximate $\varphi(x)$ for x sufficiently close to a. Now consider that φ is the exact solution of the IVP (1). Remind that this implies

$$\varphi'(x) = f(x, \varphi(x)).$$

Instead of a we take a point x_k from the partition of the interval $[x_0, x^*]$ and instead of x we take x_{k+1} which must be close to x_k. Denote, as before, an approximation of $\varphi(x_k)$ by y_k. Then

$$\varphi(x_k) + (x_{k+1} - x_k)f(x_k, \varphi(x_k))$$

is an approximation for $\varphi(x_{k+1})$. But this formula is not practical since it uses the exact value $\varphi(x_k)$ which is not known. That is why it is replaced by an approximation y_k. After this we obtain

$$y_{k+1} = y_k + (x_{k+1} - x_k)f(x_k, y_k).$$

The second idea uses the geometrical interpretation of a differential equation, more exactly the notion of *direction field*. We will see that these directions are tangent to the solution curves of the differential equations and that, an approximate solution is constructed "following" these directions as close as possible. Since the direction field is an important tool also in the qualitative methods, we will present this notion together with some examples.

The *direction field, also called slope field* in $\mathbb{R}^2$ of the scalar differential equation $y' = f(x, y)$ (with $f : \mathbb{R}^2 \to \mathbb{R}$ a continuous function) is a collection of vectors. For an arbitrary given point $(x, y) \in \mathbb{R}^2$, such a vector is based in (x, y) and have the slope $m = f(x, y)$. This number $m = f(x, y)$ it is said to be the slope of the direction field in the point (x, y).

For example, considering the differential equation

$$y' = 1 - \frac{x}{y^2} \tag{2}$$

the slope of its direction field in the point $(0, 1)$ is 1, that means that the corresponding vector in $(0, 1)$ is parallel to the first bisectrix. Also, the slope of its direction field in the point $(1, 1)$ is 0, that means that the corresponding vector in $(1, 1)$ is parallel to the Ox-axis. Although the right hand side of the equation is not defined in the point $(1, 0)$, we say that the slope in $(1, 0)$ is ∞, that means that the corresponding vector in $(1, 0)$ is parallel to the Oy-axis.

In order to have a clearer picture of the direction field it is useful to "organize" the vectors finding some isoclines. The *isocline for the slope* m is the curve

$$I_m = \{(x, y) \,:\, f(x, y) = m\}.$$

For example, for (2), the isocline for the slope 1 is the curve of equation

$$1 - \frac{x}{y^2} = 1,$$

that after simplification gives the line

$$x = 0.$$

Also, the isocline for the slope 0 is the parabola

$$y^2 = x.$$

The usefulness of the direction field comes after the following property. *The slope of the direction field in some given point is the slope of the solution curve that passes through that point.* More precisely, let the point (x_1, y_1) be given and let a solution $\varphi(x)$ of $y' = f(x, y)$, whose graph passes through this point. We know that the slope of the direction field is $f(x_1, y_1)$ and that the slope of the solution curve is $\varphi'(x_1)$. We have to prove that

$$\varphi'(x_1) = f(x_1, y_1).$$

Indeed, since the graph of φ passes through (x_1, y_1) we have that $\varphi(x_1) = y_1$, and since φ is a solution of $y' = f(x, y)$ we have that $\varphi'(x_1) = f(x_1, \varphi(x_1))$. The proof is done.

Now we come back to the Euler numerical method to find an approximate solution of the IVP $y' = f(x, y),\ y(x_0) = y_0$. The geometrical idea behind it is the following. We start in (x_0, y_0) and follow the vector of slope $f(x_0, y_0)$ until it intersects the vertical line $x = x_1$ in a point (x_1, y_1). Remind that x_0, y_0, x_1 are given, and deduce that y_1 satisfies

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

Thus

$$y_1 = y_0 + (x - x_0)f(x_0, y_0).$$

Once we are in (x_1, y_1) we follow the vector of slope $f(x_1, y_1)$ until it intersects the vertical line $x = x_2$ in a point (x_2, y_2) with

$$y_2 = y_1 + f(x_1, y_1)(x_2 - x_1).$$

We proceed in the same way until the end of the interval, x^*, obtaining

$$y_{k+1} = y_k + f(x_k, y_k)(x_{k+1} - x_k).$$

Now we continue our study of the direction field with a second example, where we consider the differential equation

$$y' = -\frac{x}{y}.$$

We will find the shape of the solution curves using the direction field. First we notice that, given m, the isocline for the slope m is the line

$$y = -\frac{1}{m}x.$$

We notice that the vectors of the directions field are orthogonal to the corresponding isocline. Hence, any solution curve is orthogonal to all the lines that passes through the origin of coordinates. We deduce that a solution curve must be a circle centered at the origin.

The *direction field* in the phase space $\mathbb{R}^2$ of a planar dynamical system is defined in a similar way and we will see that it is tangent to its orbits. More precisely, let $f_1, f_2 : \mathbb{R}^2 \to \mathbb{R}$ be continuous functions and let

$$\dot{x} = f_1(x, y), \quad \dot{y} = f_2(x, y).$$

By definition, the slope of the direction field in the point (x, y) is

$$m = \frac{f_2(x, y)}{f_1(x, y)}.$$

We have the following useful property. *The slope of the direction field in some given point is the slope of the orbit that passes through that point.* Indeed, let the point (x_1, y_1) be given and let a solution $(\varphi_1(t), \varphi_2(t))$ of the system whose orbit passes through this point. We know that the vector $(\varphi_1'(t), \varphi_2'(t))$ is tangent to the orbit for any t. Take now t_1 such that $(\varphi_1(t_1), \varphi_2(t_1)) = (x_1, y_1)$. Note that $(\varphi_1'(t_1), \varphi_2'(t_1)) = (f_1(x_1, y_1), f_2(x_1, y_1))$ and that this vector is tangent to the orbit in (x_1, y_1). Hence the slope of the orbit in (x_1, y_1) is $f_2(x_1, y_1)/f_1(x_1, y_1)$. The proof is done.