Math53: Ordinary Differential Equations
Autumn 2004

Unit 5 Summary
Numerical Methods

Extremely Important: first derivative and first-order approximation of a function.

Very Important: first- and second-order numerical methods; second-order approximation of a function.

Important: higher-order numerical methods; error estimates; why numerical methods might fail.

Description

(1) The numerical methods discussed in class are used to estimate the value $y(b)$ of the solution $y=y(t)$ to a *first-order* IVP

$$y' = f(t,y), \qquad y(a) = y_0, \tag{1}$$

at b, for $b>a$. These methods apply to IVPs involving a system of *first-order* ODEs,

$$\mathbf{y}' = \mathbf{f}(t,\mathbf{y}), \qquad \mathbf{y}(a) = \mathbf{y}_0,$$

just as well. Since any ODE, or a system of ODEs, can be written as a system of first-order ODEs, these methods can also be applied to higher-order equations and systems, but indirectly. Numerical methods are especially useful when the explicit solution $y=y(t)$, or $\mathbf{y}=\mathbf{y}(t)$, cannot be found. They can also be used if f, or $\mathbf{f}$, is known only at a discreet grid of points (t_i, y_j), or $(t_i, \mathbf{y}_j)$, as the case may well be in an experimental setting.

(2) All numerical methods encountered in this course are *fixed-step* methods. This means that we break up the interval $[a,b]$ into N segments $[t_i, t_{i+1}]$ of equal length $h=(b-a)/N$, i.e.

$$t_0 = a, \quad t_1 = t_0+h = a+h, \ \ldots \ t_{N-1} = t_{N-2}+h = a+(N-1)h, \quad t = t_{N-1}+h = a+Nh = b.$$

We then give an estimate y_i for the value of the function y at t_i. More precisely, we give an estimate y_1 for $y(t_1)$, where $y=y(t)$ is the solution to the IVP

$$y' = f(t,y), \qquad y(t_0) = y_0.$$

We then use the same procedure to give an estimate y_2 for $\tilde{y}_1(t_2)$, where $\tilde{y}_1=\tilde{y}_1(t)$ is the solution to the IVP

$$y' = f(t,y), \qquad y(t_1) = y_1.$$

Since y_1 is an estimate for $y(t_1)$, y_2 will also be an estimate for $y(t_2)$. At the ith step of this construction, we give an estimate y_{i+1} for $\tilde{y}_i(t_{i+1})$, where $\tilde{y}_i=\tilde{y}_i(t)$ is the solution to the IVP

$$y' = f(t,y), \qquad y(t_i) = y_i.$$

After N steps, we end up with an estimate y_N for $\tilde{y}_{N-1}(t_N)$, where $\tilde{y}_{N-1}=\tilde{y}_{N-1}(t)$ is the solution to the IVP

$$y' = f(t,y), \qquad y(t_N) = y_{N-1}.$$

This number y_N will also be an estimate for $y(b)$.

(3) The general mechanism described in (2) applies to all numerical methods of Chapter 6. What distinguishes them is how the estimate y_{i+1} for $\tilde{y}_i(t_{i+1})$ is obtained from (t_i, y_i) and f. In the simplest, *first-order* or Euler's, method, we take

$$y_{i+1} = y_i + s_i h, \qquad \text{where} \qquad s_i = f(t_i, y_i).$$

Since $\tilde{y}_i(t_i)=y_i$ and $\tilde{y}_i'(t_i)=f(t_i,y_i)=s_i$, y_{i+1} estimates $\tilde{y}_i(t_{i+1})$ to first order in h. In particular, this first-order method ignores the second and higher-order derivatives of $\tilde{y}_i$ at t_i. In the *second-order* Runge-Kutta, or improved Euler's, method, we take

$$y_{i+1} = y_i + \frac{s_{i,1}+s_{i,2}}{2}h, \qquad \text{where} \qquad s_{i,1} = f(t_i, y_i), \quad s_{i,2} = f(t_{i+1}, y_i + s_{i,1}h).$$

In other words, we take into account not only $\tilde{y}_i'(t_i)$, but also the change in the derivative of $\tilde{y}_i$. The above expression for y_{i+1} can be written as

$$\begin{gathered} y_{i+1}=y_i + \frac{s_{i,1}+s_{i,2}}{2}h = y_i + s_{i,1}h + \frac{1}{2}\cdot\frac{s_{i,2}-s_{i,1}}{h}\cdot h^2, \\ \text{where} \qquad s_{i,1}=f(t_i,y_i), \quad s_{i,2}=f(t_{i+1}, y_i+s_{i,1}h). \end{gathered}$$

Since $s_i=\tilde{y}_i'(t_i)$ and $s_{i+1}\approx\tilde{y}_i'(t_{i+1})$, $(s_{i,2}-s_{i,1})/h$ is an estimate for $\tilde{y}_i''(t)$. Thus, this second-order method takes into account the first and second derivatives of $\tilde{y}_i$, but ignores the third and higher derivatives of $\tilde{y}_i$. On the other hand, the *fourth-order* Runge-Kutta takes into account the first four derivatives by considering the slopes at even more points.

Error Estimates

(1) In the first-order numerical method, we ignore the second and higher-order derivatives. Thus, the error at each step is bounded by a multiple of h^2:

$$\big|\tilde{y}_i(t_{i+1}) - y_{i+1}\big| \le A_f h^2.$$

Since the number N of steps is proportional to h^{-1}, the total error in this case is bounded by a multiple of h:

$$\big|y(b) - y_N\big| = \big|y(t_N) - y_N\big| \le \tilde{A}_f h.$$

The constants A_f and $\tilde{A}_f$ depend on the length of the interval (a,b) and on the maximum values of the functions $|f(t,y)|$, $|f_t(t,y)|$, and $|f_y(t,y)|$, for $t\in[a,b]$ and $y\in(-\infty,\infty)$, as described in the book and you derive in PS7-Problem E. However, if one can find constants c and d such that it can be shown that $c<y_i<d$ for all $i=1,\ldots,N$ and $c<y(t)<d$ for all $t\in[a,b]$, where $y=y(t)$ is the solution to (eq1), the above constants A_f and $\tilde{A}_f$ will depend on the maximum values of the functions $|f(t,y)|$, $|f_t(t,y)|$, and $|f_y(t,y)|$, for $t\in[a,b]$ and only for $y\in(c,d)$.

(2) In the second-order numerical method, we ignore the third and higher-order derivatives. Thus, the error at each step is bounded by a multiple of h^3:

$$\big|\tilde{y}_i(t_{i+1}) - y_{i+1}\big| \le B_f h^3.$$

Since the number N of steps is proportional to h^{-1}, the total error in this case is bounded by a multiple of h^2:

$$\big|y(b) - y_N\big| = \big|y(t_N) - y_N\big| \le \tilde{B}_f h^2.$$

The constants B_f and $\tilde{B}_f$ depend on the length of the interval (a,b) and on the maximum values of the norms of $f(t,y)$ and of its first and second partial derivatives, with respect to t and y, for $t\in[a,b]$ and $y\in(-\infty,\infty)$. Finally, in the fourth-order numerical method, we ignore the fifth and higher-order derivatives, and thus

$$\big|\tilde{y}_i(t_{i+1}) - y_{i+1}\big| \le C_f h^5 \quad\Longrightarrow\quad \big|y(b) - y_N\big| = \big|y(t_N) - y_N\big| \le \tilde{C}_f h^4.$$

The constants C_f and $\tilde{C}_f$ depend on the length of the interval (a,b) and on the maximum values of the norms of $f(t,y)$ and of its first–fourth partial derivatives, with respect to t and y, for $t\in[a,b]$ and $y\in(-\infty,\infty)$. While the second- and fourth-order methods involve more complicated steps, they require far fewer steps to get the answer with a desired precision. If we would like to estimate $y(b)$ with an error no larger than ϵ, the time required is about $C_1\epsilon^{-1}$ for the first-order method and about $C_4\epsilon^{-1/4}$ for the fourth-order method, provided the round-off errors are insignificant. The positive constants C_1 and C_4 do not depend on ϵ, but may be very different. However, if ϵ is very small, $C_4\epsilon^{-1/4}$ is *much* smaller than $C_1\epsilon^{-1}$.

(3) There are two types of cases when fixed-step numerical methods fail. For example, the solution to the initial value problem

$$y' = y^2, \qquad y(1) = 1,$$

is $y(t)=1/(2-t)$, for $t\in(-\infty,2)$. *Please check this by solving this IVP and directly.* This solution $y=y(t)$ is not defined for $t\ge 2$. However, each of the above numerical methods will produce a finite estimate y_N for $y(b)$, for any b. If $b\ge 2$, y_N will increase rapidly as the step size h drops, and the number N of steps increases. This rapid increase will suggest that $y(b)$ is not defined, but in order to see that this is happening, various step sizes have to be tried. This example does not contradict the error bounds in (1) and (2) above, because $f_y=2y$ has no maximum, and thus there are no error bounds.

(4) Another possibility of failure is illustrated in Section 6.5. The solution to the initial value problem

$$y' = f(t,y) = t(y-1), \qquad y(-10) = 1,$$

is $y(t) = 1-e^{(t^2-100)/2}$, for $t\in(-\infty,\infty)$. *Please check this by solving this IVP and directly.* If $b\ge 10$, $y(b)\le 0$. On the other hand, if t is close 0, e.g. $|t|<1$, $y(t)$ is *extremely* close to 1. So is each estimate y_i for $y(t_i)$ if t_i is close to 1. Thus, it is very likely that for some i, the estimate y_i will be assigned the value 1, possibly because of a round-off error. If so, y_j will be 1 for all $j>i$, since $f(t,1)=0$ for all t. In such a case, our estimate y_N for $y(b)$ will be 1, unless the step size is *extremely* small. This happens because the constants in the above error estimates depend exponentially on the length of the interval $[a,b]$. Furthermore, even if h is taken to be extremely small, we may still end up with $y_N=1$ for $y(b)$, for all $b\ge 0$, due to round-off errors.