MA 214: Introduction to numerical analysis
Lecture 28

Shripad M. Garge.
IIT Bombay

(shripad@math.iitb.ac.in)

2021-2022

Improper integrals

Improper integrals result when either the function is unbounded or the interval is unbounded.

The latter case is where the interval is (a, ∞), $(-\infty, b)$ or $(-\infty, \infty)$.

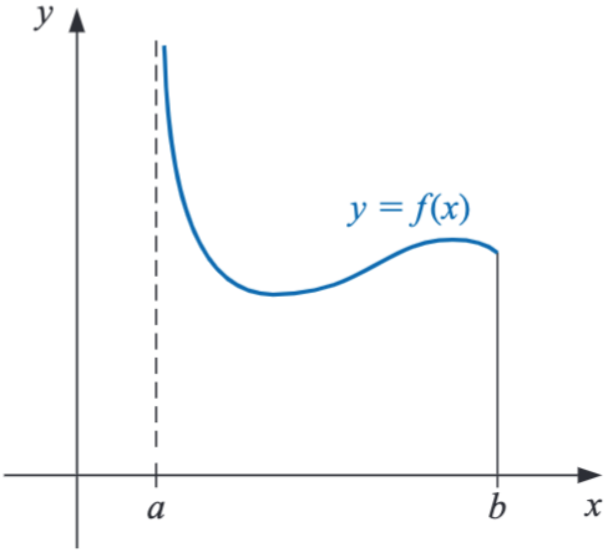
In either circumstance, the normal rules of integral approximation must be modified.

We will first consider the situation when the integrand is unbounded at the left endpoint of the interval of integration.

In this case we say that f has a singularity at the endpoint a.

In fact, other improper integrals can be reduced to problems of this form so this is a general case.

Improper integrals



Improper integrals

Our calculus courses tell us that the improper integral with a singularity at the left endpoint

$$\int_a^b \frac{dx}{(x-1)^p}$$

converges if and only if $0 < p < 1$, and in this case we define

$$\int_a^b \frac{dx}{(x-a)^p} = \lim_{M\to a^+} \frac{(x-a)^{1-p}}{1-p}\Bigg|_{x=M}^{x=b} = \frac{(b-a)^{1-p}}{1-p}.$$

For instance, the improper integral $\displaystyle\int_0^1 \frac{dx}{\sqrt{x}}$ converges (to 2) but $\displaystyle\int_0^1 \frac{dx}{x^2}$ diverges.

If $f(x) = \dfrac{g(x)}{(x-a)^p}$ where $0 < p < 1$ and $g : [a, b] \to \mathbb{R}$ is continuous then the improper integral

$$\int_a^b f(x)dx$$

exists.

We will approximate this integral using the composite Simpson's rule, provided that g is 5-times continuously differentiable on $[a, b]$, to be able to use the fourth Taylor polynomial, $P_4(x)$, for g about a:

$$P_4(x) = g(a) + g'(a)(x-a) + \frac{g''(a)}{2}(x-a)^2 + \frac{g'''(a)}{3!}(x-a)^3 + \frac{g^{(4)}(a)}{4!}(x-a)^4.$$

Improper integrals

We write

$$\int_a^b f(x)dx = \int_a^b \frac{g(x) - P_4(x)}{(x-a)^p}dx + \int_a^b \frac{P_4(x)}{(x-a)^p}dx.$$

The latter integral can be computed easily, as P_4 is a polynomial, which is

$$\sum_{k=0}^{4} \int_a^b \frac{g^{(k)}(a)}{k!}(x-a)^{k-p}dx = \sum_{k=0}^{4} \frac{g^{(k)}(a)}{k!(k+1-p)}(b-a)^{k+1-p}.$$

This is generally the dominant portion of the approximation, especially when the $P_4(x)$ agrees closely with $g(x)$ throughout the interval $[a, b]$.

We now need to determine $\displaystyle\int_a^b \frac{g(x) - P_4(x)}{(x-a)^p}\,dx$. We define

$$G(x) = \begin{cases} \frac{g(x)-P_4(x)}{(x-a)^p} & \text{if } a < x \leqslant b \\ 0 & \text{if } x = a. \end{cases}$$

Since $0 < p < 1$ and $P_4^{(k)}(a)$ agrees with $g^{(k)}(a)$ for $k = 0, \ldots, 4$, G is 4-times continuously differentiable on $[a, b]$.

Therefore composite Simpson's rule can be applied to approximate the integral of G on $[a, b]$.

Adding this approximation to the value obtained in the previous slide gives an approximation to the improper integral of f on $[a, b]$, within the accuracy of the composite Simpson's rule.

Compute $\displaystyle\int_0^1 \frac{e^x}{\sqrt{x}}dx$.

Here $P_4(x) = 1 + x + \dfrac{x^2}{2} + \dfrac{x^3}{6} + \dfrac{x^4}{24}$, so

$$\begin{aligned}
\int_0^1 \frac{P_4(x)}{\sqrt{x}}dx &= \int_0^1 \left(x^{-1/2} + x^{1/2} + \frac{1}{2}x^{3/2} + \frac{1}{6}x^{5/2} + \frac{1}{24}x^{7/2}\right) dx \\
&= \lim_{M\to 0^+} \left[2x^{1/2} + \frac{2}{3}x^{3/2} + \frac{1}{5}x^{5/2} + \frac{1}{21}x^{7/2} + \frac{1}{108}x^{9/2}\right]_M^1 \\
&= 2 + \frac{2}{3} + \frac{1}{5} + \frac{1}{21} + \frac{1}{108} \\
&\approx 2.92354497354.
\end{aligned}$$

The example, continued

Now we need to approximate $\int_0^1 G(x)dx$ where

$$G(x) = \begin{cases} \frac{1}{\sqrt{x}}(e^x - P_4(x)) & \text{if } 0 < x \leqslant 1 \\ 0 & \text{if } x = 0. \end{cases}$$

Using composite Simpson's rule, we get

$$\begin{aligned} \int_0^1 G(x)dx &\approx \frac{(0.25)}{3}[G(0) + 4G(0.25) + 2G(0.5) + 4G(0.75) + G(1)] \\ &= \frac{(0.25)}{3}[0 + 4(0.0000170) + 2(0.0004013) \\ &\quad +4(0.0026026) + 0.0099485] = 0.0017691. \end{aligned}$$

Hence

$$\int_0^1 \frac{e^x}{\sqrt{x}}dx \approx 2.9235450 + 0.0017691 = 2.9253141.$$

Singularity at b

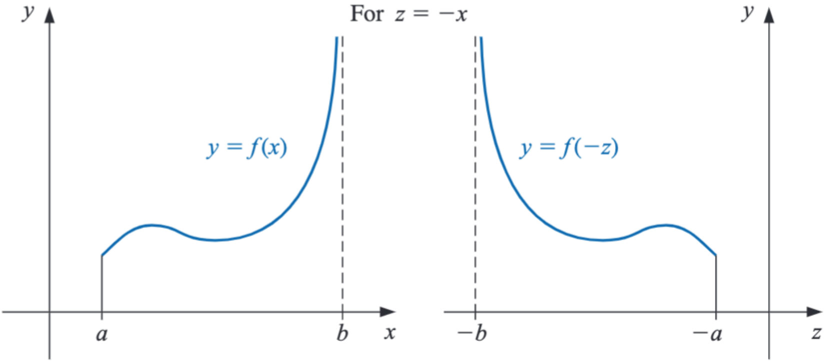
If the integrand has a singularity at the right endpoint, b, and not at a, then we could develop a similar technique making appropriate changes.

Alternatively, we can make the substitution $z = -x$, $dz = -dx$ to get

$$\int_a^b f(x)dx = \int_{-b}^{-a} f(-z)dz$$

which has a singularity at the left endpoint now.

Then we can apply the left endpoint singularity technique we have already developed.


y
For $z = -x$
y
$y = f(x)$
$y = f(-z)$
a
b
x
$-b$
$-a$
z

Improper integrals

If the integrand f has singularities at both the end points or if it has a singularity at $c \in (a, b)$ then we can write

$$\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx$$

and use the techniques developed above.

The other type of improper integral involves infinite limits of integration.

The basic integral of this type has the form $\displaystyle\int_a^\infty \frac{1}{x^p}dx$ for $p > 1$.

We use the substitution

$$t = x^{-1}, \quad dt = -x^{-2}dx, \quad \text{so} \quad dx = -x^2 dt = -t^{-2}dt.$$

Then

$$\int_a^\infty \frac{1}{x^p}dx = \int_{1/a}^0 -\frac{t^p}{t^2}dt = \int_0^{1/a} \frac{1}{t^{2-p}}dt.$$

Similarly, if f is defined on $[a,\infty)$ then we use the substitution $t = x^{-1}$ to obtain

$$\int_a^\infty f(x)dx = \int_0^{1/a} t^{-2} f\left(\frac{1}{t}\right) dt.$$

It can now be approximated by the techniques developed above.

If $a = 0$ or if both the endpoints are infinite then we write

$$\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx$$

for a suitable $a < c < b$ and apply the above substitutions.

The fourth theme is complete

This completes the fourth theme of our course:

Numerical Differentiation and Integration.

From our next lecture, we will begin the next theme:

Ordinary Differential Equations.

MA 214: Introduction to numerical analysis
Lecture 29

Shripad M. Garge.
IIT Bombay

(shripad@math.iitb.ac.in)

2021-2022

Ordinary differential equations

Differential equations are used to model problems in science and engineering that involve the change of some variable with respect to another.

Most of these problems require the solution of an initial-value problem, that is, the solution to a differential equation that satisfies a given initial condition.

In common real-life situations, the differential equation that models the problem is too complicated to solve exactly, and one of two approaches is taken to approximate the solution.

The first approach is to modify the problem by simplifying the differential equation to one that can be solved exactly and then use the solution of the simplified equation to approximate the solution to the original problem.

Ordinary differential equations

The other approach, which we will examine here, uses methods for approximating the solution of the original problem.

This is the approach that is most commonly taken because the approximation methods give more accurate results and realistic error information.

The methods that we consider here do not produce a continuous approximation to the solution of the initial-value problem.

Rather, approximations are found at certain specified, and often equally spaced, points.

Some basic material

A function $f(t,y)$ is said to satisfy a Lipschitz condition in the variable y on a set $D \subset \mathbb{R}^2$ if a constant $L > 0$ exists with

$$|f(t,y_1) - f(t,y_2)| \leqslant L|y_1 - y_2|$$

whenever (t,y_1) and (t,y_2) are in D. The constant L is called a Lipschitz constant for f.

Theorem

Let $D = \{(t,y) : a \leqslant t \leqslant b, -\infty < y < \infty\}$ and $f(t,y)$ be continuous on D. If f satisfies a Lipschitz condition on D in the variable y, then the initial-value problem

$$y'(t) = f(t,y), \quad y(a) = \alpha$$

has a unique solution $y(t)$ for $t \in [a,b]$.

Well-posed problems

Initial-value problems obtained by observing physical phenomena generally only approximate the true situation, so we need to know whether small changes in the statement of the problem introduce correspondingly small changes in the solution.

This is also important because of the introduction of round-off error when numerical methods are used.

We say that an initial-value problem is well-posed if

(i) it has a unique solution,   and

(ii) an initial-value problem obtained by small perturbations also has a unique solution.

We make the above part more precise.

Well-posed problem

Consider the initial-value problem

$$\frac{dy}{dt} = f(t, y), \;\; a \leqslant t \leqslant b, \qquad y(a) = \alpha.$$

We say that this is a well-posed problem if

- a unique solution, $y(t)$, to the problem exists, and

- there are $\epsilon_0 > 0$, $k > 0$ such that for any $\epsilon_0 > \epsilon > 0$, whenever $\delta(t)$ is continuous with $|\delta(t)| < \epsilon$ over $[a, b]$ and when $\delta_0 < \epsilon$, the initial-value problem

$$\frac{dz}{dt} = f(t, z) + \delta(t), \;\; a \leqslant t \leqslant b, \qquad z(a) = \alpha + \delta_0$$

has a unique solution, $z(t)$, that satisfies

$$|z(t) - y(t)| < k\epsilon \quad \text{for all } t \in [a, b].$$

Perturbed problem

The problem specified above, whose solution is $z(t)$, is called a perturbed problem associated with the original problem.

It assumes the possibility of an error being introduced in the statement of the differential equation, as well as an error δ_0 being present in the initial condition.

Numerical methods will always be concerned with solving a perturbed problem because any round-off error introduced in the representation perturbs the original problem.

Unless the original problem is well-posed, there is little reason to expect that the numerical solution to a perturbed problem will accurately approximate the solution to the original problem.

Condition ensuring well-posed property

The following theorem specifies conditions that ensure that an initial-value problem is well-posed.

Theorem

Suppose $D = \{(t, y) | a \leqslant t \leqslant b, -\infty < y < \infty\}$. If f is continuous and satisfies a Lipschitz condition in the variable y on the set D, then the initial-value problem

$$\frac{dy}{dt} = f(t, y), \quad a \leqslant t \leqslant b, \qquad y(a) = \alpha$$

is well-posed.

Now, we study the numerical methods to solve the initial-value problems.

Euler's method

Euler's method is the most elementary approximation technique for solving initial-value problems.

Although it is seldom used in practice, the simplicity of its derivation can be used to illustrate the techniques involved in the construction of some of the more advanced techniques.

The object of Euler's method is to obtain approximations to the well-posed initial-value problem

$$\frac{dy}{dt} = f(t, y), \quad a \leqslant t \leqslant b, \qquad y(a) = \alpha.$$

We will not obtain a continuous solution but we generate approximations to y at various values, called mesh points, in the interval $[a, b]$.

Euler's method

Once the approximate solution is obtained at the points, the approximate solution at other points in the interval can be found by interpolation.

We select the mesh points $t_i = a + ih$, for each $i = 0, 1, \ldots, N$ for a positive integer N.

The common distance between the points

$$h = (b - a)/N = t_{i+1} - t_i$$

is called the step size.

We will use Taylor's Theorem to derive Euler's method.

Euler's method

Suppose that $y(t)$, the unique solution to the problem, has two continuous derivatives on $[a,b]$, so that for each $i = 0, \ldots, N-1$,

$$\begin{aligned} y(t_{i+1}) &= y(t_i) + (t_{i+1} - t_i)y'(t_i) + \frac{(t_{i+1} - t_i)^2}{2}y''(\xi_i) \\ &= y(t_i) + hy'(t_i) + \frac{h^2}{2}y''(\xi_i) \end{aligned}$$

for some $\xi_i \in (t_i, t_{i+1})$. This then gives

$$y(t_{i+1}) = y(t_i) + h\ f(t_i, y(t_i)) + \frac{h^2}{2}y''(\xi_i).$$

Euler's method constructs $w_i \approx y(t_i)$, for each i by deleting the remainder term.

We take $w_0 = \alpha$ and $w_{i+1} = w_i + hf(t_i, w_i)$ for $i \geqslant 0$.

An example

We illustrate Euler's method by means of an example.

Consider the initial-value problem

$$y' = y - t^2 + 1, \;\; 0 \leqslant t \leqslant 2, \qquad y(0) = 0.5.$$

Let us suppose that we want to approximate $y(2)$.

We take $h = 0.5$. Here $f(t, y) = y - t^2 + 1$. We have $w_0 = 0.5$,

$$w_1 = w_0 + (0.5)(w_0 - (0.0)^2 + 1) = 0.5 + 0.5(1.5) = 1.25,$$
$$w_2 = 2.25, \;\; w_3 = 3.375, \qquad \text{and} \qquad y(2) \approx w_4 = 4.4375.$$

Geometrically, when w_i is a close approximation to $y(t_i)$, the well-posed property implies that

$$f(t_i, w_i) \approx f(t_i, y(t_i)) = y'(t_i).$$

We now take $N = 10$ and compare the results with the actual values given by

$$y(t) = (t+1)^2 - 0.5e^t.$$

We then have $h = 0.2$, $t_i = (0.2)i$, $w_0 = 0.5$ and

$$\begin{aligned} w_{i+1} &= w_i + h(w_i - t_i^2 + 1) \\ &= w_i + (0.2)[w_i - 0.04i^2 + 1] \\ &= 1.2w_i - 0.008i^2 + 0.2 \end{aligned}$$

for $0 \leqslant i \leqslant 9$. This gives $w_1 = 0.8$, $w_2 = 1.152$ and so on.

We compare the approximate values and the actual values in a table in the following slide.

The example, continued

t_i	w_i	$y_i = y(t_i)$	$\lvert y_i - w_i \rvert$
0.0	0.5000000	0.5000000	0.0000000
0.2	0.8000000	0.8292986	0.0292986
0.4	1.1520000	1.2140877	0.0620877
0.6	1.5504000	1.6489406	0.0985406
0.8	1.9884800	2.1272295	0.1387495
1.0	2.4581760	2.6408591	0.1826831
1.2	2.9498112	3.1799415	0.2301303
1.4	3.4517734	3.7324000	0.2806266
1.6	3.9501281	4.2834838	0.3333557
1.8	4.4281538	4.8151763	0.3870225
2.0	4.8657845	5.3054720	0.4396874

The example, continued

When we had $h = 0.5$, the approximate value for $y(2)$ was 4.4375 whereas for $h = 0.2$ it is 4.8657845, much closer to the actual value.

Note that the error grows slightly as the value of t increases.

This controlled error growth is a consequence of the stability of Euler's method, which implies that the error is expected to grow in no worse than a linear manner.

We will study the error analysis of Euler's method in our next lecture.

MA 214: Introduction to numerical analysis
Lecture 30

Shripad M. Garge.
IIT Bombay

(shripad@math.iitb.ac.in)

2021-2022

Euler's method

We consider an initial-value problem

$$\frac{dy}{dt} = f(t, y), \quad a \leqslant t \leqslant b, \qquad y(a) = \alpha$$

and give an approximate solution at equally spaced nodes in $[a, b]$.

The method begins by choosing a positive integer N, letting $h = \frac{b-a}{N}$ and taking the nodes, $t_i = a + ih$ for $i = 0, \ldots, N$.

We then define $w_0 = \alpha$ and $w_{i+1} = w_i + hf(t_i, w_i)$ for $i \geqslant 0$.

We did an example in the last lecture with two different values of N.

Now, we study the error pattern in this method.

Error in Euler's method

Theorem

Suppose f is continuous and satisfies a Lipschitz condition with the constant L on $D = [a, b] \times \mathbb{R}$.

Assume that the unique solution, $y(t)$, to the initial-value problem

$$\frac{dy}{dt} = f(t, y), \;\; a \leqslant t \leqslant b, \qquad y(a) = \alpha$$

satisfies $|y''(t)| \leqslant M$.

If w_i for $i = 0, \ldots, N$, are the Euler approximations to $y(t_i)$ then

$$|y(t_i) - w_i| \leqslant \frac{hM}{2L}\left[e^{L(t_i - a)} - 1\right].$$

Error in Euler's method

The result is true if $i = 0$. Let us denote $y(t_i)$ by y_i. Then

$$y_{i+1} - w_{i+1} = (y_i - w_i) + h[f(t_i, y_i) - f(t_i, w_i)] + \frac{h^2}{2} y''(\xi_i)$$

and

$$\begin{aligned}
|y_{i+1} - w_{i+1}| &\leqslant (1 + hL)|y_i - w_i| + \frac{h^2 M}{2} \\
&\leqslant (1 + hL)\left((1 + hL)|y_{i-1} - w_{i-1}| + \frac{h^2 M}{2}\right) + \frac{h^2 M}{2} \\
&\leqslant e^{(i+1)hL}\left(|y_0 - w_0| + \frac{h^2 M}{2hL}\right) - \frac{h^2 M}{2hL} \\
&\leqslant \frac{hM}{2L}\left(e^{(t_{i+1}-a)L} - 1\right).
\end{aligned}$$

Error in Euler's method

The weakness of the above theorem is that we need to know a bound for $y''(t)$.

Although this condition often prohibits us from obtaining a realistic error bound, it should be noted that if $\partial f/\partial t$ and $\partial f/\partial y$ both exist, the chain rule gives

$$y''(t) = \frac{dy'}{dt}(t) = \frac{df}{dt}(t, y(t)) = \frac{\partial f}{\partial t}(t, y(t)) + \frac{\partial f}{\partial y}(t, y(t)) \cdot f(t, y(t)).$$

So it is, at times, possible to obtain an error bound for $y''(t)$ without explicitly knowing $y(t)$.

The principal importance of the error formula is that the bound depends linearly on the step size h. Consequently, diminishing the step size gives greater accuracy to the approximations.

We did not consider the round-off error while deriving the above formula.

As h becomes smaller, more calculations are necessary and more round-off error is expected.

In actuality then, we do not use

$$w_0 = \alpha, \qquad w_{i+1} = w_i + hf(t_i, w_i)$$

but

$$u_0 = \alpha + \delta_0, \qquad u_{i+1} = u_i + hf(t_i, u_i) + \delta_{i+1}$$

where $\delta_0, \ldots$ are the round-off errors.

We now describe an error bound for Euler's method taking the round-off errors into account.

Error in Euler's method

Consider the initial-value problem

$$\frac{dy}{dt} = f(t,y), \quad a \leqslant t \leqslant b, \qquad y(a) = \alpha$$

satisfying the hypothesis of the above theorem.

Let $u_0, \ldots$ be the approximations in the Euler's method with round-off errors δ_i and assume that $|\delta_i| < \delta$ for all i.

Then

$$|y(t_i) - u_i| \leqslant \frac{1}{L}\left(\frac{hM}{2} + \frac{\delta}{h}\right)\left[e^{L(t_i - a)} - 1\right] + |\delta_0| e^{L(t_i - a)}$$

Note that this error bound is not linear in h. In fact, the error will be expected to be large for small step size.

Since the object of a numerical technique is to determine accurate approximations with minimal effort, we need a means for comparing the efficiency of various approximation methods.

The first device we consider is called the local truncation error of the method.

The local truncation error at a specified step measures the amount by which the exact solution to the differential equation fails to satisfy the difference equation being used for the approximation at that step.

We want to know how well the approximations generated by the method satisfy the differential equation, not the other way around.

However, we don't know the exact solution so we cannot generally determine this.

Local truncation error

The local truncation serves quite well to determine not only the local error of a method but also the actual approximation error.

Consider the initial-value problem

$$y' = f(t, y), \quad a \leqslant t \leqslant b, \qquad y(a) = \alpha.$$

The difference method $w_0 = \alpha$, $w_{i+1} = w_i + h\phi(t_i, w_i)$ for $i = 0, \ldots, N-1$, has local truncation error

$$\tau_{i+1}(h) = \frac{y_{i+1} - (y_i + h\phi(t_i, y_i))}{h} = \frac{y_{i+1} - y_i}{h} - \phi(t_i, y_i).$$

This error is a local error because it measures the accuracy of the method at a specific step, assuming that the method was exact at the previous step.

Local truncation error

Consider Euler's method, for instance, where the difference equation is $w_0 = \alpha$ and $w_{i+1} = w_i + hf(t_i, w_i)$.

The local truncation error is

$$\tau_{i+1}(h) = \frac{y_{i+1} - y_i}{h} - f(t_i, y_i) = \frac{h}{2} y''(\xi_i).$$

If one knows that $|y''(t)| \leqslant M$ for some M then the local truncation error is bounded by $\dfrac{hM}{2}$.

So the local truncation error in Euler's method is $O(h)$.

We select difference-equation methods for solving initial-value problems in such a manner that their local truncation errors are $O(h^p)$ for as large a value of p as possible, while keeping the number and complexity of calculations within a reasonable bound.

Higher order Taylor methods

Since Euler method used Taylor's theorem, we try using Taylor's theorem for higher orders.

Consider the initial-value problem

$$y' = f(t, y), \;\; a \leqslant t \leqslant b, \qquad y(a) = \alpha$$

with f being n-times continuously differentiable.

By Taylor's theorem

$$y(t_{i+1}) = y(t_i) + hy'(t_i) + \cdots + \frac{h^n}{n!}y^{(n)}(t_i) + \frac{h^{(n+1)}}{(n+1)!}y^{(n+1)}(\xi_i)$$

for some $\xi \in (t_i, t_{i+1})$.

We can compute the derivatives of y at t_i using the function $f(t, y)$.

Higher order Taylor methods

Successive differentiation of $y(t)$ gives

$$y'(t) = f(t, y(t)), \;\; y''(t) = f'(t, y(t)), \cdots, y^{(n+1)}(t) = f^{(n)}(t, y(t)).$$

Putting this in the above formula, we get

$$y_{i+1} = y_i + hf(t_i, y_i) + \cdots + \frac{h^n}{n!} f^{(n-1)}(t_i, y_i) + \frac{h^{(n+1)}}{(n+1)!} f^{(n)}(\xi_i, y(\xi_i)).$$

The Taylor method of order n is then $w_0 = \alpha$ and

$$w_{i+1} = w_i + hf(t_i, w_i) + \cdots + \frac{h^n}{n!} f^{(n-1)}(t_i, w_i).$$

Euler's method is Taylor method of order one.

MA 214: Introduction to numerical analysis
Lecture 31

Shripad M. Garge.
IIT Bombay

(shripad@math.iitb.ac.in)

2021-2022

Taylor method of order n

Consider the initial-value problem

$$\frac{dy}{dt} = f(t, y), \quad a \leqslant t \leqslant b, \qquad y(a) = \alpha.$$

The Taylor method of order n gives an approximate solution to the above problem at equally spaced nodes $t_i = a + ih$ where $h = \frac{b-a}{N}$.

These solutions are $w_0 = \alpha$ and $w_{i+1} = w_i + hT(t_i, w_i)$ where

$$T(t_i, w_i) = f(t_i, w_i) + \frac{h}{2} f'(t_i, w_i) + \cdots + \frac{h^{n-1}}{n!} f^{(n-1)}(t_i, w_i).$$

Let us now do some examples.

An example

Let us apply Taylor's method of order two with $N = 10$ to solve

$$y' = y - t^2 + 1, \ \ 0 \leqslant t \leqslant 2, \qquad y(0) = 0.5.$$

We will need to compute $y'' = f'(t, y)$:

$$f'(t, y(t)) = \frac{d}{dt}(y - t^2 + 1) = y' - 2t = (y - t^2 + 1) - 2t = y - t^2 - 2t + 1.$$

Then

$$\begin{aligned} T(t_i, w_i) &= f(t_i, w_i) + \frac{h}{2} f'(t_i, w_i) \\ &= w_i - t_i^2 + 1 + \frac{h}{2}(w_i - t_i^2 - 2t_i + 1) \\ &= \left(1 + \frac{h}{2}\right)(w_i - t_i^2 + 1) - ht_i. \end{aligned}$$

The example, continued

Since $N = 10$, we get that $h = 0.2$ and $t_i = 0.2i$ for $i = 0, \dots, 10$. Then the method gives $w_0 = 0.5$ and

$$
\begin{aligned}
w_{i+1} &= w_i + h\left[\left(1 + \frac{h}{2}\right)(w_i - t_i^2 + 1) - ht_i\right] \\
&= w_i + (0.2)\left[\left(1 + \frac{0.2}{2}\right)(w_i - 0.04i^2 + 1) - 0.04i\right] \\
&= 1.22w_i - 0.0088i^2 - 0.008i + 0.22.
\end{aligned}
$$

We have

$$y(0.2) \approx 0.83 \quad \text{and} \quad y(0.4) \approx 1.2158.$$

We now list the values and the errors in the following table.

The example, continued

Order 2 and $N = 10$.

t_i	w_i	$\lvert y(t_i) - w_i \rvert$
0.0	0.500000	0
0.2	0.830000	0.000701
0.4	1.215800	0.001712
0.6	1.652076	0.003135
0.8	2.132333	0.005103
1.0	2.648646	0.007787
1.2	3.191348	0.011407
1.4	3.748645	0.016245
1.6	4.306146	0.022663
1.8	4.846299	0.031122
2.0	5.347684	0.042212

Another example

We now apply Taylor's method of order 4 with $N = 10$ to the same problem

$$y' = y - t^2 + 1, \;\; 0 \leqslant t \leqslant 2, \qquad y(0) = 0.5.$$

We compute

$$\begin{aligned}
f'(t, y(t)) &= y - t^2 - 2t + 1, \\
f''(t, y(t)) &= \frac{d}{dt}(y - t^2 - 2t + 1) = y' - 2t - 2 \\
&= (y - t^2 + 1) - 2t - 2 = y - t^2 - 2t - 1, \\
f'''(t, y(t)) &= \frac{d}{dt}(y - t^2 - 2t - 1) = y' - 2t - 2 \\
&= y - t^2 - 2t - 1.
\end{aligned}$$

The example, continued

Then

$$
\begin{aligned}
T(t_i, w_i) &= f(t_i, w_i) + \frac{h}{2} f'(t_i, w_i) + \frac{h^2}{6} f''(t_i, w_i) + \frac{h^3}{24} f'''(t_i, w_i) \\
&= (w_i - t_i^2 + 1) + \frac{h}{2}(w_i - t_i^2 - 2t_i + 1) \\
&\quad + \frac{h^2}{6}(w_i - t_i^2 - 2t_i - 1) + \frac{h^3}{24}(w_i - t_i^2 - 2t_i - 1) \\
&= \left(1 + \frac{h}{2} + \frac{h^2}{6} + \frac{h^3}{24}\right)(w_i - t_i^2) - \left(1 + \frac{h}{3} + \frac{h^2}{12}\right)(ht_i) \\
&\quad + 1 + \frac{h}{2} - \frac{h^2}{6} - \frac{h^3}{24}.
\end{aligned}
$$

The method then gives $w_0 = 0.5$ and

$$\begin{aligned} w_{i+1} &= w_i + hT(t_i, w_i) \\ &= w_i + h\left[\left(1 + \frac{h}{2} + \frac{h^2}{6} + \frac{h^3}{24}\right)(w_i - t_i^2)\right. \\ &\quad \left. - \left(1 + \frac{h}{3} + \frac{h^2}{12}\right)(ht_i) + 1 + \frac{h}{2} - \frac{h^2}{6} - \frac{h^3}{24}\right] \end{aligned}$$

With $N = 10$ and $h = 0.2$, this becomes

$$w_{i+1} = 1.2214\ w_i - 0.008856\ i^2 - 0.00856\ i + 0.2186.$$

$$y(0.2) \approx 0.8293 \quad \text{and} \quad y(0.4) \approx 1.214091.$$

We note the results in the next table.

The example, continued

Order 4 and $N = 10$.

t_i	w_i	$\lvert y(t_i) - w_i \rvert$
0.0	0.500000	0
0.2	0.829300	0.000001
0.4	1.214091	0.000003
0.6	1.648947	0.000006
0.8	2.127240	0.000010
1.0	2.640874	0.000015
1.2	3.179964	0.000023
1.4	3.732432	0.000032
1.6	4.283529	0.000045
1.8	4.815238	0.000062
2.0	5.305555	0.000083

The example, continued

t_i	w_i	$\lvert y(t_i) - w_i \rvert$
0.0	0.500000	0
0.2	0.830000	0.000701
0.4	1.215800	0.001712
0.6	1.652076	0.003135
0.8	2.132333	0.005103
1.0	2.648646	0.007787
1.2	3.191348	0.011407
1.4	3.748645	0.016245
1.6	4.306146	0.022663
1.8	4.846299	0.031122
2.0	5.347684	0.042212

Order 2

t_i	w_i	$\lvert y(t_i) - w_i \rvert$
0.0	0.500000	0
0.2	0.829300	0.000001
0.4	1.214091	0.000003
0.6	1.648947	0.000006
0.8	2.127240	0.000010
1.0	2.640874	0.000015
1.2	3.179964	0.000023
1.4	3.732432	0.000032
1.6	4.283529	0.000045
1.8	4.815238	0.000062
2.0	5.305555	0.000083

Order 4

The example, continued

Suppose now that with this data, we want to approximate the value of y at some other point, say at $y = 1.25$.

We then use interpolation.

If we use linear interpolation on the nodes 1.2 and 1.4 then we get

$$\begin{aligned} y(1.25) &\approx \left(\frac{1.25-1.4}{1.2-1.4}\right) 3.1799640 + \left(\frac{1.25-1.2}{1.4-1.2}\right) 3.7324321 \\ &= 3.3180810. \end{aligned}$$

The true value is $y(1.25) = 3.3173285$, so the error of this approximation is 0.0007525, about 30 times the average of the errors at 1.2 and 1.4.

We can improve upon this by using cubic Hermite interpolation on the nodes 1.2 and 1.4.

The example, continued

For that we will need to compute the values of y' at these two nodes:

$$y'(1.2) = y(1.2) - (1.2)^2 + 1 \approx 2.7399640$$

and

$$y'(1.4) = y(1.4) - (1.4)^2 + 1 \approx 2.7724321.$$

The corresponding cubic Hermite polynomial is

$$3.1799640 + (t - 1.2)(2.7399640) + (t - 1.2)^2(0.1118825)$$
$$+(t - 1.2)^2(t - 1.4)(-0.3071225)$$

Then

$$\begin{aligned} y(1.25) &\approx 3.1799640 + 0.1369982 + 0.0002797 + 0.0001152 \\ &= 3.3173571. \end{aligned}$$

The example, continued

This value is accurate to within 0.0000286, which is about the average of the errors at 1.2 and at 1.4, and only 4% of the error obtained using linear interpolation.

This improvement in the accuracy certainly justifies the added computations required for the Hermite method.

Error in Taylor's method

Theorem

If Taylor's method of order n is used to approximate the solution to

$$y' = f(t,y), \;\; a \leqslant t \leqslant b, \qquad y(a) = \alpha$$

with step size h then the local truncation error is $O(h^n)$.

There are, of course, limitations and draw backs to this method.

The number of calculations of various derivatives of $f(t, y(t))$ is a bit of a concern, even if the error is $O(h^n)$.

We would like to reduce the number of calculations while keeping the error to the similar order.

MA 214: Introduction to numerical analysis
Lecture 32

Shripad M. Garge.
IIT Bombay

(shripad@math.iitb.ac.in)

2021-2022

Taylor method of order n

The Taylor methods outlined in the previous lectures have the desirable property of high-order local truncation error.

But they also have the disadvantage of requiring the computation and evaluation of the derivatives of $f(t, y)$.

This is a complicated and time-consuming procedure for most problems, so the Taylor methods are seldom used in practice.

We would like to learn approximation methods to solve the initial-value problems which do not require as high computations and evaluations but have the error to the similar order.

We will study the Runge-Kutta method but before that we will need the statement of Taylor's theorem in two variables.

Taylor's theorem in two variables

Theorem

Suppose that $f(t, y)$ and all its partial derivatives of order less than or equal to $n+1$ are continuous on

$$D = \{(t, y) : a \leqslant t \leqslant b, c \leqslant y \leqslant d\},$$

and let $(t_0, y_0) \in D$. For every $(t, y) \in D$, there exists ξ between t and t_0 and μ between y and y_0 with

$$f(t, y) = P_n(t, y) + R_n(t, y)$$

where P_n is called the n-th Taylor polynomial in two variables for the function f about (t_0, y_0), and $R_n(t, y)$ is the remainder term associated with $P_n(t, y)$.

The error term $R_n(t, y)$ involves ξ and μ.

Taylor's theorem in two variables

$$
\begin{aligned}
P_n(t,y) \quad = \quad & f(t_0,y_0) + \left[(t-t_0)\frac{\partial f}{\partial t}(t_0,y_0) + (y-y_0)\frac{\partial f}{\partial y}(t_0,y_0)\right] \\
& + \left[\frac{(t-t_0)^2}{2}\frac{\partial^2 f}{\partial t^2}(t_0,y_0) + (t-t_0)(y-y_0)\frac{\partial^2 f}{\partial t \partial y}(t_0,y_0) + \right. \\
& \left. \frac{(y-y_0)^2}{2}\frac{\partial^2 f}{\partial y^2}(t_0,y_0)\right] + \cdots \\
& + \left[\frac{1}{n!}\sum_{j=0}^{n}\binom{n}{j}(t-t_0)^{n-j}(y-y_0)^j \frac{\partial^n f}{\partial t^{n-j}\partial y^j}(t_0,y_0)\right]
\end{aligned}
$$

and

$$
R_n(t,y) = \frac{1}{(n+1)!}\sum_{j=0}^{n+1}\binom{n+1}{j}(t-t_0)^{n+1-j}(y-y_0)^j \frac{\partial^{n+1} f}{\partial t^{n+1-j}\partial y^j}(\xi,\mu).
$$

Runge-Kutta Methods of Order Two

The first step in deriving a Runge-Kutta method is to determine values for a_1, α_1 and β_1 with the property that $a_1 f(t+\alpha_1, y+\beta_1)$ approximates

$$T(t,y) = f(t,y) + \frac{h}{2} f'(t,y)$$

with error no greater than $O(h^2)$. Since

$$\begin{aligned} f'(t,y) &= \frac{df}{dt}(t,y) = \frac{\partial f}{\partial t}(t,y) + \frac{\partial f}{\partial y}(t,y) \cdot y'(t) \\ &= \frac{\partial f}{\partial t}(t,y) + \frac{\partial f}{\partial y}(t,y) \cdot f(t,y) \end{aligned}$$

we get

$$T(t,y) = f(t,y) + \frac{h}{2}\frac{\partial f}{\partial t}(t,y) + \frac{h}{2}\frac{\partial f}{\partial y}(t,y) \cdot f(t,y).$$

Runge-Kutta Methods of Order Two

Using Taylor's theorem of order 1 we get

$$\begin{aligned} a_1 f(t+\alpha_1, y+\beta_1) &= a_1 f(t,y) \\ &+ a_1\alpha_1 \frac{\partial f}{\partial t}(t,y) + a_1\beta_1 \frac{\partial f}{\partial y}(t,y) \\ &+ a_1 R_1(t+\alpha_1, y+\beta_1) \end{aligned}$$

where

$$R_1(t+\alpha_1, y+\beta_1) = \frac{\alpha_1^2}{2}\frac{\partial^2 f}{\partial t^2}(\xi,\mu) + \alpha_1\beta_1 \frac{\partial^2 f}{\partial t \partial y}(\xi,\mu) + \frac{\beta_1^2}{2}\frac{\partial^2 f}{\partial y^2}(\xi,\mu)$$

for some ξ between t and $t+\alpha_1$ and μ between y and $y+\beta_1$.

We now match the equations for $T(t,y)$ and $a_1 f(t+\alpha_1, y+\beta_1)$ and get some equations.

Runge-Kutta Methods of Order Two

These equations are

$$f(t,y): a_1 = 1, \qquad \frac{\partial f}{\partial t}(t,y): a_1\alpha_1 = \frac{h}{2}, \qquad \frac{\partial f}{\partial y}(t,y): a_1\beta_1 = \frac{h}{2}f(t,y).$$

We then conclude

$$a_1 = 1, \qquad \alpha_1 = \frac{h}{2}, \qquad \text{and} \qquad \beta_1 = \frac{h}{2}f(t,y)$$

and get

$$T(t,y) = f\left(t + \frac{h}{2}, y + \frac{h}{2}f(t,y)\right) - R_1\left(t + \frac{h}{2}, y + \frac{h}{2}f(t,y)\right)$$

where the error R_1 can be written as follows:

$$
\begin{aligned}
R_1\left(t+\frac{h}{2}, y+\frac{h}{2} f(t, y)\right) &= \frac{h^2}{8} \frac{\partial^2 f}{\partial t^2}(\xi, \mu) \\
&+\frac{h^2}{4} f(t, y) \frac{\partial^2 f}{\partial t \partial y}(\xi, \mu) \\
&+\frac{h^2}{8} \frac{\partial^2 f}{\partial y^2}(\xi, \mu)
\end{aligned}
$$

If all the second-order partial derivatives of f are bounded, then

$$
R_1\left(t+\frac{h}{2}, y+\frac{h}{2} f(t, y)\right)
$$

is of $O(h^2)$. As a consequence, the order of error for this method is the same as that of the Taylor method of order two.

Runge-Kutta Methods of Order Two

The difference-equation method resulting from replacing $T(t, y)$ in Taylor's method of order two by

$$f\left(t + \frac{h}{2}, y + \frac{h}{2} f(t, y)\right)$$

is a specific Runge-Kutta method known as the Midpoint-method.

It gives $w_0 = \alpha$ and

$$w_{i+1} = w_i + hf\left(t_i + \frac{h}{2}, w_i + \frac{h}{2} f(t_i, w_i)\right)$$

for $i = 0, \ldots, N - 1$.

We do an example now.

Use the Midpoint method with $N = 10$, $h = 0.2$, $t_i = 0.2i$, and $w_0 = 0.5$ to approximate the solution to our usual example

$$y' = y - t^2 + 1, \ \ 0 \leqslant t \leqslant 2, \qquad y(0) = 0.5.$$

The difference formula is

$$w_{i+1} = 1.22w_i - 0.0088i^2 - 0.008i + 0.218.$$

for each $i = 0, \ldots, 9$.

The first two steps of these methods give

$$w_1 = 0.828, \qquad \text{and} \qquad w_2 = 1.21136.$$

The table is given on the next slide.

t_i	$y(t_i)$	Midpoint Method	Error
0.0	0.5000000	0.5000000	0
0.2	0.8292986	0.8280000	0.0012986
0.4	1.2140877	1.2113600	0.0027277
0.6	1.6489406	1.6446592	0.0042814
0.8	2.1272295	2.1212842	0.0059453
1.0	2.6408591	2.6331668	0.0076923
1.2	3.1799415	3.1704634	0.0094781
1.4	3.7324000	3.7211654	0.0112346
1.6	4.2834838	4.2706218	0.0128620
1.8	4.8151763	4.8009586	0.0142177
2.0	5.3054720	5.2903695	0.0151025