MA 214: Introduction to numerical analysis
Lecture 21

Shripad M. Garge.
IIT Bombay

(shripad@math.iitb.ac.in)

2021-2022

Three-point endpoint formula:

$$f'(x_0) = \frac{1}{2h}\big(-3f(x_0) + 4f(x_0+h) - f(x_0+2h)\big) + \frac{h^2}{3}f^{(3)}(\xi_0)$$

where ξ_0 lies between x_0 and $x_0 + 2h$.

Three-point midpoint formula:

$$f'(x_0) = \frac{1}{2h}\big(f(x_0+h) - f(x_0-h)\big) - \frac{h^2}{6}f^{(3)}(\xi_0)$$

where ξ_0 lies between $x_0 - h$ and $x_0 + h$.

Both the errors are of the order of h^2 but the second error is about half of the first error.

Three-point formula

We recall that in the initial formula, the error was of the order of h, here it is of the order of h^2.

It is therefore clear that if we apply this method for $f(x) = \ln x$ then the results will be much better.

The endpoint formula for $f(x) = \ln x$, $x_0 = 1.8$ and $h = 0.1$ gives the approximation 0.55454184711, much better approximation than the initial method for $h = 0.01$.

The midpoint formula for the same f, x_0 and h gives approximation 0.55612817555 which is closer to the actual value of $f'(x_0)$.

We now proceed to the five-point formulae.

Five-point formula

Let f be five times continuously differentiable on an interval I and let us assume that the nodes in the following formulae are in I.

Five-point midpoint formula:

$$\begin{aligned} f'(x_0) &= \frac{1}{12h}\big[f(x_0-2h) - 8f(x_0-h) + 8f(x_0+h) - f(x_0+2h)\big] \\ &+ \frac{h^4}{30}f^{(5)}(\xi) \quad \text{for some } \xi \text{ between } x_0-2h \text{ and } x_0+2h. \end{aligned}$$

Five-point endpoint formula:

$$\begin{aligned} f'(x_0) &= \frac{1}{12h}\big[-25f(x_0) + 48f(x_0+h) - 36f(x_0+2h) \\ &+ 16f(x_0+3h) - 3f(x_0+4h)\big] + \frac{h^4}{5}f^{(5)}(\xi) \end{aligned}$$

where ξ lies between x_0 and x_0+4h.

Midpoint vis-à-vis endpoint formulae

The midpoint formulae are simpler to operate than the endpoint formulae. The errors in the midpoint formulae are also better than the ones in the endpoint formulae.

However, the point x_0 may be closer to one of the endpoints of the interval I and some of the nodes, in the midpoint formula, may not lie in I.

In such a case, instead of taking h very small, we will use the endpoint formula.

Note that the endpoint formula is applicable for $h < 0$ also, while the midpoint formula stays the same if we change the sign of h.

What next?

One can also derive formulae for computing derivatives of higher order.

We will see one such formula, just to get an idea but we will not discuss them in detail.

$$f''(x_0) = \frac{1}{h^2}\big[f(x_0 - h) - 2f(x_0) + f(x_0 + h)\big] - \frac{h^2}{12} f^{(4)}(\xi)$$

where ξ is between $x_0 - h$ and $x_0 + h$.

This is the midpoint formula, as you would have guessed.

Round-off error

It is now time to discuss our old friend (and nemesis), the round-off error!

Our three-point midpoint formula is

$$f'(x_0) = \frac{1}{2h}\big(f(x_0+h) - f(x_0-h)\big) - \frac{h^2}{6} f^{(3)}(\xi_0).$$

Suppose that, in evaluating $f(x_0 \pm h)$ we encounter round-off errors $e(x_0 \pm h)$, respectively.

Then our computations actually use the values $\tilde{f}(x_0+h)$ and $\tilde{f}(x_0-h)$, which are related to the true values by

$$f(x_0 \pm h) = \tilde{f}(x_0 \pm h) + e(x_0 \pm h).$$

Round-off error

The total error in the approximation is then

$$\frac{e(x_0+h)-e(x_0-h)}{2h} - \frac{h^2}{6} f^{(3)}(\xi_0).$$

If we assume that the round-off errors $e(x_0 \pm h)$ are bounded by some number $\epsilon > 0$ and that $f^{(3)}$ is bounded by a number $M > 0$, then

$$\left| \frac{e(x_0+h)-e(x_0-h)}{2h} - \frac{h^2}{6} f^{(3)}(\xi_0) \right| \leqslant \frac{\epsilon}{h} + \frac{h^2}{6} M.$$

To reduce the truncation error, $h^2 M/6$, we should reduce h. But then the round-off error, ϵ/h, grows. In practice, then, it is seldom advantageous to let h be too small, because then the round-off error will dominate the calculations.

An example

Consider the following values of $\sin x$:

x	$\sin x$	x	$\sin x$
0.800	0.71736	0.901	0.78395
0.850	0.75128	0.902	0.78457
0.880	0.77074	0.905	0.78643
0.890	0.77707	0.910	0.78950
0.895	0.78021	0.920	0.79560
0.898	0.78208	0.950	0.81342
0.899	0.78270	1.000	0.84147

We approximate $f'(0.900)$, where $f(x) = \sin x$, using

$$f'(0.900) \approx \frac{f(0.900 + h) - f(0.900 - h)}{2h}$$

for different values of h.

The example continued

Here are the results:

h	approximation to $f'(0.900)$	error
0.001	0.62500	0.00339
0.002	0.62250	0.00089
0.005	0.62200	0.00039
0.010	0.62150	-0.00011
0.020	0.62150	-0.00011
0.050	0.62140	-0.00021
0.100	0.62055	-0.00106

The errors have been computed using $\cos(0.900) = 0.62161$. The optimal choice for h appears to lie between 0.005 and 0.05.

The example continued

When the error is

$$e(h) = \frac{\epsilon}{h} + \frac{h^2}{6}M$$

one uses calculus to prove that the minimum value of $e(h)$ occurs at

$$h = \sqrt[3]{3\epsilon/M}.$$

Since

$$M = \max_{[0.8,1]} |f'''(x)| = \max_{[0.8,1]} |\cos x| = \cos 0.8 \approx 0.69671.$$

The values of f are given to 5 decimal places, hence the round-off error is 5×10^{-6} and

$$h = \sqrt[3]{\frac{3 \times 0.000005}{0.69671}} \approx 0.02781931326.$$

Round-off error

In practice, we cannot compute an optimal h while approximating the derivative, since we have no knowledge of the third derivative of the function.

But we must remain aware that reducing the step size will not always improve the approximation.

This is why we use the endpoint formula when x_0 is close to one of the endpoints of the interval I.

We have considered the round-off error problems for the three-point formula, but similar difficulties occur with all the differentiation formulas.

The reason can be traced to the need to divide by a power of h.

Instability

We understand it well that division by small numbers tends to exaggerate the round-off error, and this operation should be avoided if possible.

In the case of numerical differentiation, we cannot avoid the problem entirely, although the higher-order methods reduce the difficulty.

As an approximation method, numerical differentiation is unstable, since the small values of h needed to reduce truncation error also cause the round-off error to grow.

This is the first class of unstable methods we have encountered, and these techniques need to be avoided if possible.

However, in addition to being used for computational purposes, the formulae are needed for approximating the solutions of ordinary and partial differential equations.

MA 214: Introduction to numerical analysis
Lecture 22

Shripad M. Garge.
IIT Bombay

(shripad@math.iitb.ac.in)

2021-2022

Numerical integration

The need often arises for evaluating the definite integral of a function that has no explicit antiderivative or whose antiderivative is not easy to obtain. The basic method involved in approximating

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

is called numerical quadrature.

It uses a sum $\sum_{i=0}^{n} a_i f(x_i)$ to approximate $\int_a^b f(x)dx$.

It goes without saying that interpolating polynomials will be immensely useful here too.

We select distinct nodes $\{x_0, \ldots, x_n\}$ from the interval $[a, b]$. Then integrate the Lagrange interpolating polynomial

$$P_n(x) = \sum_{i=0}^{n} f(x_i) L_i(x)$$

and its truncation error term over $[a, b]$ to obtain

$$\begin{aligned}\int_a^b f(x)dx &= \int_a^b \sum_{i=0}^{n} f(x_i) L_i(x) dx + \int_a^b \prod_{i=0}^{n} (x - x_i) \frac{f^{(n+1)}(\xi(x))}{(n+1)!} dx \\ &= \sum_{i=0}^{n} a_i f(x_i) + \frac{1}{(n+1)!} \int_a^b \prod_{i=0}^{n} (x - x_i) f^{(n+1)}(\xi(x)) dx\end{aligned}$$

where $\xi(x)$ is in $[a, b]$ for each x and $a_i = \int_a^b L_i(x) dx$.

Numerical quadrature

We then say that

$$\int_a^b f(x)dx \approx \sum_{i=0}^{n} a_i f(x_i)$$

and the error is

$$E(f) = \frac{1}{(n+1)!}\int_a^b \prod_{i=0}^{n}(x - x_i) f^{(n+1)}(\xi(x))dx.$$

We now proceed by taking equally spaced nodes in $[a, b]$ and apply the above formula to the interpolating polynomials.

The initial ones give us the trapezoidal and the Simpson rules.

Let $a = x_0$ and $b = x_1$. We define h to be the difference $b - a$.

Using the linear polynomial P_1 interpolating f on the nodes $x_0 < x_1$, we get

$$\int_a^b f(x)dx = \int_{x_0}^{x_1} \frac{(x - x_1)f(x_0) - (x - x_0)f(x_1)}{x_0 - x_1} dx$$

$$+\frac{1}{2}\int_{x_0}^{x_1} f''(\xi(x))(x - x_0)(x - x_1)dx.$$

By the mean value theorem for integrals, there exists $\xi \in [a, b]$ such that

$$\int_{x_0}^{x_1} f''(\xi(x))(x - x_0)(x - x_1)dx = f''(\xi)\int_{x_0}^{x_1}(x - x_0)(x - x_1)dx$$

The trapezoidal rule

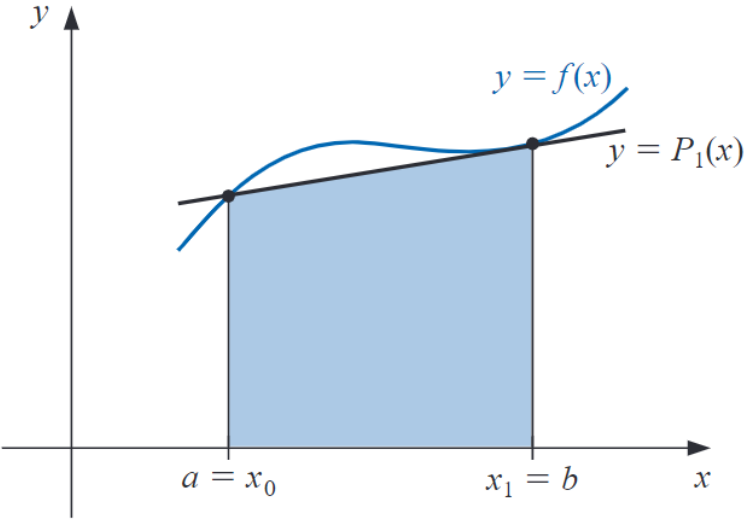
This is further equal to

$$f''(\xi)\left[\frac{x^3}{3} - \frac{(x_1 - x_0)}{2}x^2 + x_0 x_1 x\right]_{x_0}^{x_1} = -\frac{h^3}{6}f''(\xi).$$

This gives

$$\begin{aligned}
\int_a^b f(x)dx &= \left[\frac{(x-x_1)^2 f(x_0) - (x-x_0)^2 f(x_1)}{2(x_0 - x_1)}\right]_{x_0}^{x_1} - \frac{h^3}{12}f''(\xi) \\
&= \frac{(x_1 - x_0)\big(f(x_0) + f(x_1)\big)}{2} - \frac{h^3}{12}f''(\xi) \\
&= \frac{h}{2}\big(f(x_0) + f(x_1)\big) - \frac{h^3}{12}f''(\xi)
\end{aligned}$$

This is called the trapezoidal rule because when f is a function with positive values, $\int_a^b f(x)dx$ is approximated by the area in a trapezoid.



y
$y = f(x)$
$y = P_1(x)$
$a = x_0$
$x_1 = b$
x

The Simpson's rule

Simpson's rule results from using the interpolating polynomial with three equally-spaced nodes

$$x_0 = a, \quad x_1 = a + h, \quad x_2 = a + 2h = b,$$

so $h = (b - a)/2$. The rule is

$$\int_{x_0}^{x_2} f(x)dx = \frac{h}{3}[f(x_0) + 4f(x_1) + f(x_2)] - \frac{h^5}{90} f^{(4)}(\xi)$$

where ξ is an element of (x_0, x_2).

We postpone the derivation of this formula, which uses some more methods than just the interpolating polynomial P_2, to an unknown point in the future.

We instead compute a few integrals with these two formulae.

Examples

Let us consider a few functions on the interval $[0, 2]$ and use the trapezoidal as well as the Simpson's rule to compute the corresponding integrals.

We note

$$\text{Trapezoid :} \qquad \int_0^2 f(x)dx \approx f(0) + f(2)$$

$$\text{Simpson :} \quad \int_0^2 f(x)dx \approx \frac{1}{3}[f(0) + 4f(1) + f(2)]$$

For a constant function, both the formulae give the correct answer!

This should not come as a surprise because the error is given in terms of some derivative of f, which vanishes when f is a constant function.

Examples, continued

We prepare a table of integral values:

$f(x)$	Exact	Trapezoidal	Simpson
x^2	2.667	4.000	2.667
x^4	6.400	16.000	6.667
$(x+1)^{-1}$	1.099	1.333	1.111
$\sqrt{1+x^2}$	2.958	3.326	2.964
$\sin x$	1.416	0.909	1.425
e^x	6.389	8.389	6.421

MA 214: Introduction to numerical analysis
Lecture 23

Shripad M. Garge.
IIT Bombay

(shripad@math.iitb.ac.in)

2021-2022

Numerical integration

We are discussing numerical methods to compute definite integrals

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

We observed that both the trapezoidal and Simpson rules give exact answers for constant functions.

The Simpson's rule was exact even for x^2.

The degree of accuracy, or precision, of a quadrature formula is the largest positive integer n such that the formula is exact for x^k, for each $k = 0, 1, \ldots, n$.

It is an easy exercise to verify that the degrees of accuracy for the Trapezoidal and the Simpson's rules are one and three, respectively.

The Trapezoidal and Simpson's rules are examples of a class of methods known as Newton-Cotes formulae.

There are two types of Newton-Cotes formulae, closed and open.

The $(n+1)$-point closed Newton-Cotes formula uses nodes $x_i = x_0 + ih$, for $i = 0, 1, \ldots, n$, where $x_0 = a$, $x_n = b$ and $h = (b-a)/n$.

It is called 'closed' because the end points of the closed interval $[a, b]$ are included as nodes.

This formula assumes the form

$$\int_a^b f(x)dx = \sum_{i=0}^{n} a_i f(x_i) \quad \text{where} \quad a_i = \int_a^b L_i(x)dx.$$

Error formula

Let $\sum_{i=0}^{n} a_i f_i(x)$ denote the $(n+1)$-point closed Newton-Cotes formula. There exists $\xi \in (a, b)$ with

$$\int_a^b f(x)dx = \sum_{i=0}^{n} a_i f_i(x) + \frac{h^{n+3} f^{(n+2)}(\xi)}{(n+2)!} \int_0^n t^2(t-1)\cdots(t-n)dt$$

if n is even and f is $(n+2)$-times continuously differentiable, and

$$\int_a^b f(x)dx = \sum_{i=0}^{n} a_i f_i(x) + \frac{h^{n+2} f^{(n+1)}(\xi)}{(n+1)!} \int_0^n t(t-1)\cdots(t-n)dt$$

if n is odd and f is $(n+1)$-times continuously differentiable.

Note that when n is even, the degree of precision is $n+1$ even though we work with only $n+1$ nodes.

Here are some of the initial Newton-Cotes formulae:

$n = 1$: Trapezoidal rule:

$$\int_{x_0}^{x_1} f(x)dx = \frac{h}{2}\left[f(x_0) + f(x_1)\right] - \frac{h^3}{12}f''(\xi)$$

where $x_0 < \xi < x_1$.

$n = 2$: Simpson's $\frac{1}{3}$-rule:

$$\int_{x_0}^{x_2} f(x)dx = \frac{h}{3}\left[f(x_0) + 4f(x_1) + f(x_2)\right] - \frac{h^5}{90}f^{(4)}(\xi)$$

where $x_0 < \xi < x_2$.

$n = 3$: Simpson's $\frac{3}{8}$-rule:

$$\int_{x_0}^{x_3} f(x)dx = \frac{3h}{8}\left[f(x_0) + 3f(x_1) + 3f(x_2) + f(x_3)\right] - \frac{3h^5}{80} f^{(4)}(\xi)$$

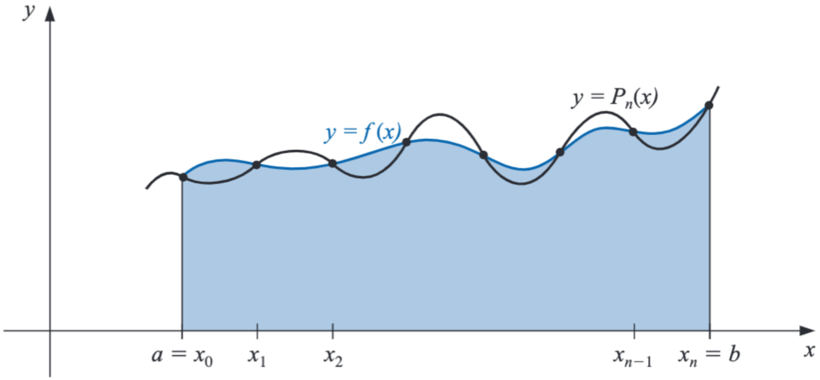
where $x_0 < \xi < x_3$.

$n = 4$:

$$\begin{aligned}\int_{x_0}^{x_4} f(x)dx &= \frac{2h}{45}\left[7f(x_0) + 32f(x_1) + 12f(x_2) + 32f(x_3) + 7f(x_4)\right] \\ &\quad - \frac{8h^7}{945} f^{(6)}(\xi)\end{aligned}$$

where $x_0 < \xi < x_4$.

Newton-Cotes formulae



Open Newton-Cotes formulae

The open Newton-Cotes formulae do not include the endpoints of $[a, b]$ as nodes. They use the nodes $x_i = x_0 + ih$, for each $i = 0, 1, \ldots, n$, where $h = (b-a)/(n+2)$ and $x_0 = a + h$.

This implies that $x_n = b - h$, so we label the endpoints by setting $x_{-1} = a$ and $x_{n+1} = b$.

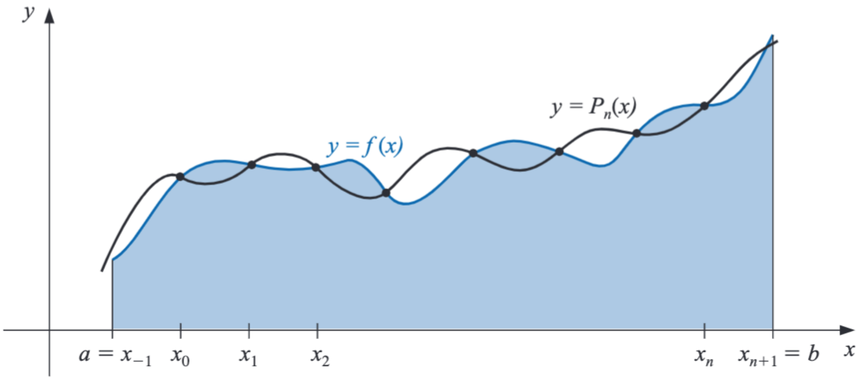
Open formulae contain all the nodes used for the approximation within the open interval (a, b). They read

$$\int_a^b f(x)dx = \int_{x_{-1}}^{x_{n+1}} f(x)dx \approx \sum_{i=0}^{n} a_i f_i(x)$$

where

$$a_i = \int_a^b L_i(x)dx.$$

Open Newton-Cotes formulae



Error formula

Let $\sum_{i=0}^{n} a_i f_i(x)$ denote the $(n+1)$-point open Newton-Cotes formula. There exists $\xi \in (a, b)$ with

$$\int_a^b f(x)dx = \sum_{i=0}^{n} a_i f_i(x) + \frac{h^{n+3} f^{(n+2)}(\xi)}{(n+2)!} \int_{-1}^{n+1} t^2(t-1)\cdots(t-n)dt$$

if n is even and f is $(n+2)$-times continuously differentiable, and

$$\int_a^b f(x)dx = \sum_{i=0}^{n} a_i f_i(x) + \frac{h^{n+2} f^{(n+1)}(\xi)}{(n+1)!} \int_{-1}^{n+1} t(t-1)\cdots(t-n)dt$$

if n is odd and f is $(n+1)$-times continuously differentiable.

Note that, as in the case of the closed methods, the degree of precision is better for the even methods than for the odd methods.

Open Newton-Cotes formulae

Some of the open Newton-Cotes formulae with their error terms are as follows:

$n = 0$: Midpoint rule:

$$\int_{x_{-1}}^{x_1} f(x)dx = 2hf(x_0) + \frac{h^3}{3} f''(\xi)$$

where $x_{-1} < \xi < x_1$.

$n = 1$:

$$\int_{x_{-1}}^{x_2} f(x)dx = \frac{3h}{2}\left[f(x_0) + f(x_1)\right] + \frac{3h^3}{4} f''(\xi)$$

where $x_{-1} < \xi < x_2$.

Open Newton-Cotes formulae

$n = 2$:

$$\int_{x_{-1}}^{x_3} f(x)dx = \frac{4h}{3}\left[2f(x_0) - f(x_1) + 2f(x_2)\right] + \frac{14h^5}{45} f^{(4)}(\xi)$$

where $x_{-1} < \xi < x_3$.

$n = 3$:

$$\int_{x_{-1}}^{x_4} f(x)dx = \frac{5h}{24}\left[11f(x_0) + f(x_1) + f(x_2) + 11f(x_3)\right] + \frac{95h^5}{144} f^{(4)}(\xi)$$

where $x_{-1} < \xi < x_4$.

Closed vis-à-vis open formulae

We compare the closed and the open formulae for the integral

$$\int_0^{\pi/4} \sin x \; dx = 1 = \frac{\sqrt{2}}{2} \approx 0.29289321881.$$

	closed	open
$n = 0$		0.30055886494
$n = 1$	0.27768018363	0.29798754218
$n = 2$	0.29293263784	0.29285865919
$n = 3$	0.29291070254	0.29286922813
$n = 4$	0.29289318256	