Babes-Bolyai University Cluj-Napoca

Faculty of Mathematics and Computer Science

Specialization: Mathematics and Computer Science

BACHELOR'S THESIS

Theme: Numerical integration methods

Kohan Alexandru

Assoc. Prof. Cătinaş Teodora

2022

Universitatea Babeş-Bolyai

Facultatea de Matematică şi Informatică

Specializare: Matematică şi Informatică

LUCRARE DE LICENŢĂ

Temă: Metode de integrare numerică

Kohan Alexandru

Assoc. Prof. Cătinaş Teodora

2022

Abstract

The main goal of this thesis is to describe certain numerical integration methods, together with their properties and solutions. There will also be some Matlab code examples as well as real-world applications of these topics. This bachelor's thesis was motivated by the practicality of numerical integration methods in our daily lives. It's used to identify irregularly shaped areas. It has a wide range of applications in physics, mathematics, chemistry, and computer science.

There are 15 chapters in the bachelor's thesis. The chapter on differentiation includes some basic information on derivatives as well as a personal example. The Richardson Extrapolation chapter presents a way to generate highly accurate results using loworder formulas. This chapter will be used at Romberg Integration. The chapter on Weighted Mean Value introduces concepts that will be applied to the Simpson and Trapezoidal Rules. The chapter Numerical integration elements presents the notion of quadrature and its elements. The Simpson and Trapezoidal rules are illustrated so that their accuracy can be compared later in a personal example. The Composite Rules chapter explains when these types of rules are beneficial and includes some personal examples. Closed Newton Cotes and Open Newton Cotes are generalizations of previously given rules, each with an example. The Romberg Integration chapter demonstrates how we may achieve high accuracy while using very little computing power. A Matlab example is also provided. The chapters on Adaptive Quadrature and Gaussian Quadrature describe two different types of quadratures, each with its own set of applications. In the last two chapters, you will find some examples of numerical integration methods in Matlab and how they could be used in real life.

Contents

1 Introduction

Most of the data of this section can be found in [1], [2], [3], [5] and [9].

The applicability of numerical integration methods in our lives were the motivating factor behind this bachelor's thesis. Petronas buildings were designed using integration methods to make them stronger. Previously, the twin towers were supposed to be only 427 meters tall. Dr Mahathir recognised the building's potential as one of the world's tallest structures and encouraged architects and engineers to add a few more meters to the height in whatever manner they could. Many structural facts had to be recalculated and retested in wind tunnels in order for him to achieve his goals. This resulted in the installation of a dome with an integrated pinnacle atop the towers, allowing it to surpass the Sears Tower in the United States by reaching 452 meters tall.

Finding volumes of wine barrels was one of the early applications of integration. One mathematician who contributed to the calculation of areas and volumes was Kepler. Kepler had invested in a wine barrel. Following that, he researched how to determine the areas and volumes of various types of entities and published a book on the subject. As a result, he made a significant contribution to the development of integral calculus. Integral calculus is now used to solve these types of problems.

2 Differentiation

Most of the results of this section can be found in [8].

The derivative of the function f at x_0 is

$$f'(x_0) = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h} \quad [8]$$

This formula demonstrates how to generate a good estimate of $f'(x_0)$; basic calculation

$$\frac{f(x_0 + h) - f(x_0)}{h}$$

for lower h values. Despite its obviousness, it completely fails because to our old enemy, the round-off error. It is, nevertheless, a good place to start. To approximate $f'(x_0)$, suppose first that $x_0 \in (a, b)$, where $f \in C^2[a, b]$, and that $x_1 = x_0 + h$ for some $h \neq 0$ that is sufficiently small to ensure that $x_1 \in [a, b]$. We construct the first Lagrange polynomial $P_{0,1}(x)$ for f determined by x_0 and x_1 , with its error term:

$$\begin{aligned} f(x) &= P_{0,1}(x) + \frac{(x - x_0)(x - x_1)}{2!} f''(\xi(x)) \\ &= \frac{f(x_0)(x - x_0 - h)}{-h} + \frac{f(x_0 + h)(x - x_0)}{h} + \\ & \frac{(x - x_0)(x - x_0 - h)}{2} f''(\xi(x)), \end{aligned}$$

for some $\xi(x)$ between x_0 and x_1. Differentiating gives

$$\begin{aligned} f'(x) =& \frac{f(x_0 + h) - f(x_0)}{h} + D_x \left[\frac{(x - x_0)(x - x_0 - h)}{2} f''(\xi(x)) \right] \\ =& \frac{f(x_0 + h) - f(x_0)}{h} + \frac{2(x - x_0) - h}{2} f''(\xi(x)) \\ &+ \frac{(x - x_0)(x - x_0 - h)}{2} D_x \left(f''(\xi(x)) \right) \end{aligned}$$

Deleting the terms involving $\xi(x)$ gives

$$f'(x) \approx \frac{f\left(x_0+h\right)-f\left(x_0\right)}{h}$$

One issue with this formula is that we don't have any data about $D_x f''(\xi(x))$, thus we can't estimate the truncation error. When x is x_0, however, the coefficient of $D_x f''(\xi(x))$ is 0, and the formula simplifies to

$$f'\left(x_0\right)=\frac{f\left(x_0+h\right)-f\left(x_0\right)}{h}-\frac{h}{2} f''(\xi)$$

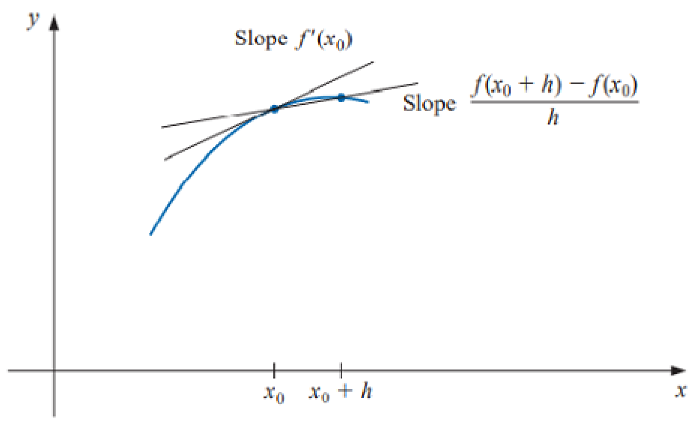
For lower h values, the difference quotient $\left[f\left(x_0+h\right)-f\left(x_0\right)\right]/h$ can be used to approximate $f'\left(x_0\right)$ with an error bounded by $M|h|/2$, where M is a bound on $|f''(x)|$ for x between x_0 and x_0+h. This formula is called the forward-difference formula if h $>$ 0 (see Figure 1) and the backward-difference formula if h $<$ 0.

Example 1 Using h $=$ 0.2, h $=$ 0.1, and h $=$ 0.02, approximate the derivative of $f(x)=1/x$ at $x_0=2.3$ using the forward-difference formula, and calculate bounds for the approximation errors.

Solution

$$\frac{f(2.3+h)-f(2.3)}{h}$$

Figure 1:



with h = 0.2 gives

$$\frac{1/2.5 - 1/2.3}{0.2} = \frac{0.4 - 0.434}{0.2} = -0.17$$

Because $f''(x) = 2/x^3$ and $2.3 < \xi < 2.5$, a bound for this approximation error is

$$\frac{|hf''(\xi)|}{2} = \frac{|h * 2|}{2\xi^3} < \frac{0.2}{2.3^3} = 0.01643$$

with h = 0.1 gives

$$\frac{1/2.4 - 1/2.3}{0.1} = \frac{0.416 - 0.434}{0.1} = -0.18$$

$$\frac{|hf''(\xi)|}{2} = \frac{|h * 2|}{2\xi^3} < \frac{0.1}{2.3^3} = 0.00821$$

with h = 0.02 gives

$$\frac{1/2.32 - 1/2.3}{0.02} = \frac{0.431 - 0.434}{0.02} = -0.15$$

$$\frac{|hf''(\xi)|}{2} = \frac{|h * 2|}{2\xi^3} < \frac{0.02}{2.3^3} = 0.00164$$

Table 1:

h	f(2.3+h)	$\frac{f(2.3+h)-f(2.3)}{h}$	$\frac{\lvert h \rvert}{2.3^3}$
0.2	0.4	-0.17	0.01643
0.1	0.416	-0.18	0.00821
0.02	0.431	-0.15	0.00164

Definition 2.1 ***Second Derivative Midpoint Formula*** [8]

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

for some ξ, where $x_0 - h < \xi < x_0 + h$.

If $f^{(4)}$ is continuous on $[x_0 - h, x_0 + h]$ it is also bounded, and the approximation is $O\left(h^2\right)$.

3 Richardson-extrapolation

Most of the results of this section can be found in [8].

When loworder formulas are used, Richardson's extrapolation is used to achieve high-accuracy results.

When it is confirmed that a technique of approximation contains an error term with a predictable form, one that is dependent on a parameter, usually the step size h, extrapolation can be used. Suppose that for each number $h \neq 0$ we have a formula $N_1(h)$ that approximates an unknown constant M, and that the approximation's truncation error has the form

$$M - N_1(h) = K_1 h + K_2 h^2 + K_3 h^3 + \cdots ,$$

for some set of (unknown) constants $K_1, K_2, K_3, \ldots$

The truncation error is O(h), so unless the magnitudes of the constants $K_1, K_2, K_3, \ldots$ differ significantly,

$$M - N_1(0.1) \approx 0.1 K_1, \quad M - N_1(0.01) \approx 0.01 K_1,$$

and, generally, $M - N_1(h) \approx K_1 h$.

Extrapolation's goal is to develop a simple approach to integrate these inaccurate O(h) approximations in a useful fashion to construct formulas with a higher-order truncation error.

Suppose, for example, we can combine the $N_1(h)$ formulas to produce an $O\left(h^2\right)$ approximation formula, $N_2(h)$, for M with

$$M - N_2(h) = \hat{K}_2 h^2 + \hat{K}_3 h^3 + \cdots ,$$

for some, again unknown, collection of constants $\hat{K}_2, \hat{K}_3, \ldots$ After that, we'd have

$$M - N_2(0.1) \approx 0.01\hat{K}_2, \quad M - N_2(0.01) \approx 0.0001\hat{K}_2$$

and so forth. If the constants K_1 and $\hat{K}_2$ are roughly of the same magnitude, then the $N_2(h)$ approximations would be much better than the corresponding $N_1(h)$ approximations. The extrapolation continues by combining the $N_2(h)$ approximations in a manner that produces formulas with $O\left(h^3\right)$ truncation error, and so on.

To see specifically how we can generate the extrapolation formulas, consider the O(h) formula for approximating M

$$M = N_1(h) + K_1 h + K_2 h^2 + K_3 h^3 + \cdots \quad (3.10)$$

The formula is assumed to hold for all positive h, so we replace the parameter h by half its value. Then we have a second O(h) approximation formula

$$M = N_1\left(\frac{h}{2}\right) + K_1\frac{h}{2} + K_2\frac{h^2}{4} + K_3\frac{h^3}{8} + \cdots \quad (3.11)$$

Subtracting Eq. (3.10) from twice Eq. (3.11) eliminates the term involving K_1 and gives

$$M = N_1\left(\frac{h}{2}\right) + \left[N_1\left(\frac{h}{2}\right) - N_1(h)\right] + K_2\left(\frac{h^2}{2} - h^2\right) + K_3\left(\frac{h^3}{4} - h^3\right) + \cdots. \quad (3.12)$$

Define

$$N_2(h) = N_1\left(\frac{h}{2}\right) + \left[N_1\left(\frac{h}{2}\right) - N_1(h)\right]$$

Then Eq. (3.12) is an $O\left(h^2\right)$ approximation formula for M:

$$M = N_2(h) - \frac{K_2}{2}h^2 - \frac{3K_3}{4}h^3 - \cdots$$

4 Weighted-mean-value

Most of the results of this section can be found in [8].

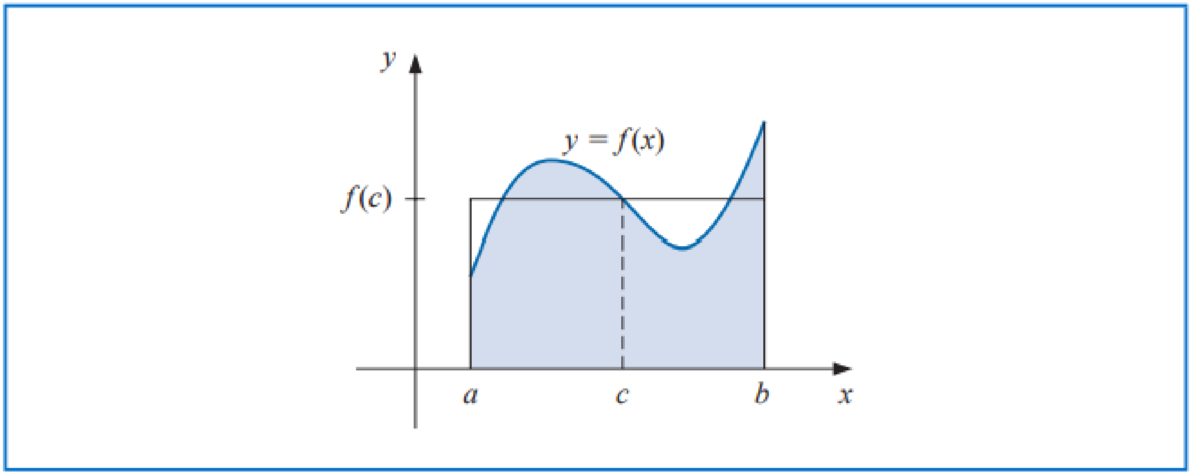
Definition 4.1 [8] Assume that $f \in C[a, b]$, that the Riemann integral of g exists on [a, b], and that g(x) does not change sign on [a, b]. Then there exists a number c in (a, b) with

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

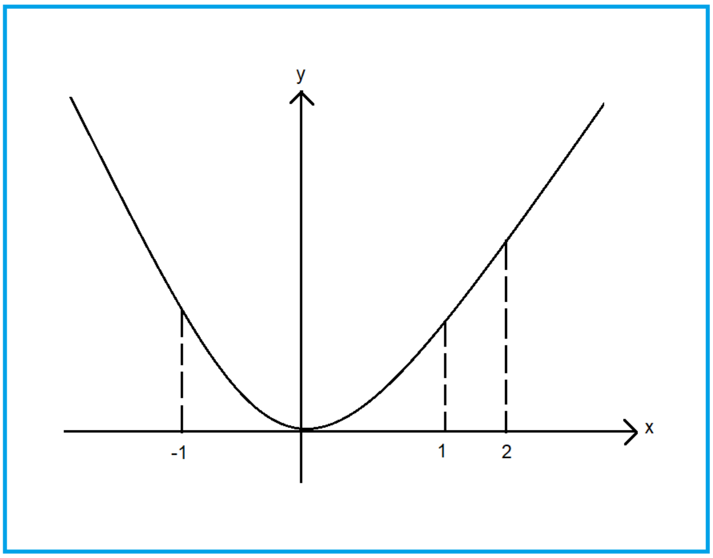
When $g(x) \equiv 1$, Theorem 4.1 is the usual Mean Value Theorem for Integrals. It calculates the average value of the function f over the interval [a, b] as (See Figure 2.)

$$f(c) = \frac{1}{b-a}\int_a^b f(x)dx$$

Figure 2:



Example: Apply Weighted Mean Value Theorem for Integrals to determine which x values the function f(x) = 1 + x^2 have the average value over the interval [-1,2]



There is a number c in [-1,2] such that

$$f(c) = \frac{1}{b-a}\int_a^b f(x)dx$$

$$h_{avg} = \frac{Area}{Width}$$

$$f(c) = f_{avg} = \frac{1}{2-(-1)}\int_{-1}^{2}(1+x^2)dx$$

$$= \frac{1}{3}\left[x + \frac{x^3}{3}\right]_{-1}^{2} = \frac{1}{3}\left[2 + \frac{8}{3} - \left(-1 - \frac{1}{3}\right)\right]$$

$$= \frac{1}{3}\left[3 + \frac{9}{3}\right] = \frac{1}{3}[3+3] = \frac{6}{3} = 2$$

$$f(c) = f_{avg} = 2$$

$$f(x) = 1 + x^2$$

$$f(c) = 2 = 1 + c^2$$

$$c^2 = 1$$

$$c = \pm\sqrt{1} = \pm 1$$

5 Numerical integration elements

Most of the results of this section can be found in [4].

Definition 5.1 (Quadrature) [4]

Formula:

$$I(f) = Q(f) + R(f)$$

where

$$Q(f) = \sum_{k=0}^{m} A_k \lambda_k(f),$$

is called a numerical integration formula (of function f) or a quadrature formula, A_k, k = 0, 1, ..., m are called the coefficients of the quadrature formula and R(f) is the remainder term

Definition 5.2 (Degree of exactness) [4]

The natural number r such that R(f) = 0, f $\in P_r$ and for which there exists g $\in P_{r+1}$ such that R(g) $\neq$ 0, is called the degree of exactness of the quadrature Q, i.e., dex(Q) = r.

Remark. [4]

The linearity of the remainder operator R implies that dex(Q) = r if and only if R(e_i) = 0, i = 0, 1, ..., r and R(e_{r+1}) $\neq$ 0, with $e_i(x) = x^i$

Consider formulas derived from first and second Lagrange polynomials with nodes spaced evenly. This results in the Trapezoidal and Simpson's rules, which are often taught in calculus classes.

6 The Trapezoidal Rule

Most of the results of this section can be found in [8].

To derive the Trapezoidal rule for approximating $\int_a^b f(x)\,dx$,

let x_0 = a, x_1 = b, h = b - a and use the linear Lagrange polynomial:

$$P_1(x) = \frac{(x-x_1)}{(x_0-x_1)} f(x_0) + \frac{(x-x_0)}{(x_1-x_0)} f(x_1).$$

$$\begin{aligned} Then \int_a^b f(x)dx = &\int_{x_0}^{x_1} \left[\frac{(x-x_1)}{(x_0-x_1)} f(x_0) + \frac{(x-x_0)}{(x_1-x_0)} f(x_1) \right] dx \\ &+ \frac{1}{2} \int_{x_0}^{x_1} f''(\xi(x))\,(x-x_0)\,(x-x_1)\,dx \quad (6.1) \end{aligned}$$

The product $(x-x_0)(x-x_1)$ does not change sign on $[x_0, x_1]$, so the Weighted Mean Value Theorem for Integrals 4.1 can be applied to the error term to give, for some ξ in (x0,x1),

$$\begin{aligned} \int_{x_0}^{x_1} f''(\xi(x))\,(x-x_0)\,(x-x_1)\,dx &= f''(\xi) \int_{x_0}^{x_1} (x-x_0)\,(x-x_1)\,dx \\ &= f''(\xi) \left[\frac{x^3}{3} - \frac{(x_1+x_0)}{2} x^2 + x_0 x_1 x \right]_{x_0}^{x_1} \\ &= -\frac{h^3}{6} f''(\xi) \end{aligned}$$

Consequently, Eq. (6.1) implies that

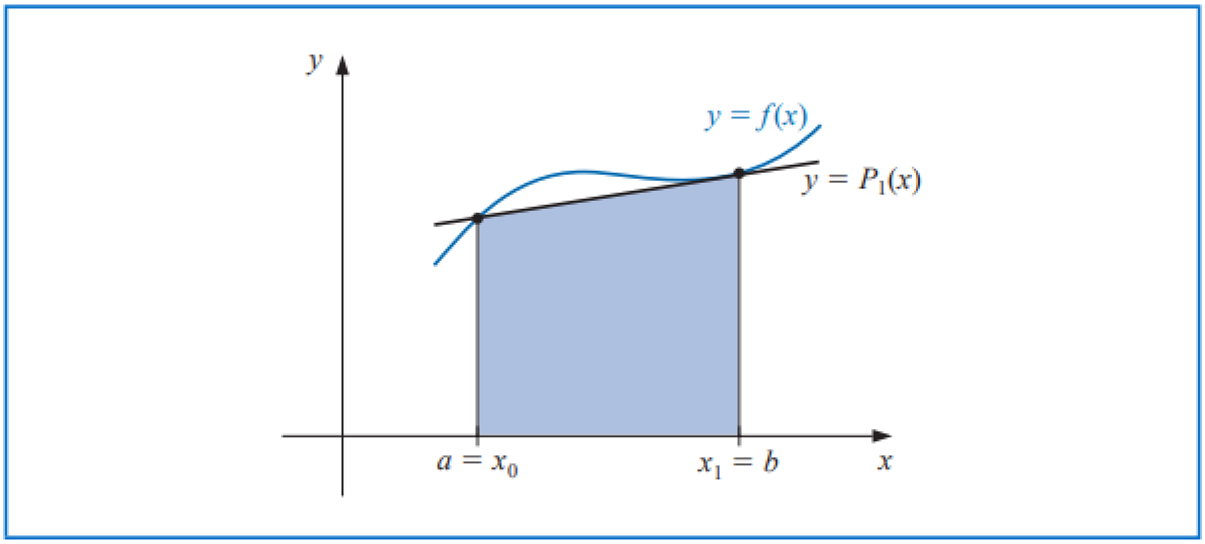
$$\begin{aligned} \int_a^b f(x)dx &= \left[\frac{(x-x_1)^2}{2\,(x_0-x_1)} f(x_0) + \frac{(x-x_0)^2}{2\,(x_1-x_0)} f(x_1) \right]_{x_0}^{x_1} - \frac{h^3}{12} f''(\xi) \\ &= \frac{(x_1-x_0)}{2} \left[f(x_0) + f(x_1) \right] - \frac{h^3}{12} f''(\xi) \end{aligned}$$

Trapezoidal Rule[8]:

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

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, as shown in Figure 3.

Figure 3:



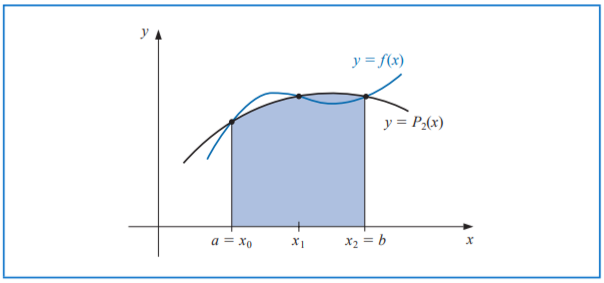
The error term for the Trapezoidal rule involves f''. When applied to any function whose second derivative is identically zero, that is, any polynomial of degree one or fewer, the Trapezoidal rule produces the precise result.

7 The Simpson Rule

Most of the results of this section can be found in [8].

Simpson's rule results from integrating over [a, b] the second Lagrange polynomial with equally-spaced nodes x_0 = a, x_2 = b, and x_1 = a + h, where h = (b - a)/2. (See Figure 4)

Figure 4:



Therefore

$$\begin{aligned}\int_a^b f(x)dx = &\int_{x_0}^{x_2}\left[\frac{(x-x_1)(x-x_2)}{(x_0-x_1)(x_0-x_2)}f(x_0)+\frac{(x-x_0)(x-x_2)}{(x_1-x_0)(x_1-x_2)}f(x_1)\right.\\ &\left.+\frac{(x-x_0)(x-x_1)}{(x_2-x_0)(x_2-x_1)}f(x_2)\right]dx\\ &+\int_{x_0}^{x_2}\frac{(x-x_0)(x-x_1)(x-x_2)}{6}f^{(3)}(\xi(x))dx.\end{aligned}$$

However, this method of deriving Simpson's rule only produces an O(h^4) error term involving $f^{(3)}$. By approaching the problem in another way, a higher-order term involving $f^{(4)}$ can be derived.

Assume that f is expanded in the third Taylor polynomial about x_1 to demonstrate this alternate way.. Then for each x in [x_0, x_2], a number ξ(x) in (x_0, x_2) exists with

$$f(x) = f(x_1) + f'(x_1)(x - x_1) + \frac{f''(x_1)}{2}(x - x_1)^2 + \frac{f'''(x_1)}{6}(x - x_1)^3 + \\ + \frac{f^{(4)}(\xi(x))}{24}(x - x_1)^4$$

and

$$\int_{x_0}^{x_2} f(x)dx = \left[f(x_1)(x - x_1) + \frac{f'(x_1)}{2}(x - x_1)^2 + \frac{f''(x_1)}{6}(x - x_1)^3 \right. \\ \left. + \frac{f'''(x_1)}{24}(x - x_1)^4 \right]_{x_0}^{x_2} + \frac{1}{24}\int_{x_0}^{x_2} f^{(4)}(\xi(x))(x - x_1)^4 \, dx. \quad (7.1)$$

Because $(x - x_1)^4$ is never negative on $[x_0, x_2]$, the Weighted Mean Value Theorem for Integrals 4.1 implies that.

$$\frac{1}{24}\int_{x_0}^{x_2} f^{(4)}(\xi(x))(x - x_1)^4 dx = \frac{f^{(4)}(\xi_1)}{24}\int_{x_0}^{x_2}(x - x_1)^4 dx = \\ = \frac{f^{(4)}(\xi_1)}{120}(x - x_1)^5]_{x_0}^{x_2},$$

for some number ξ_1 in (x_0, x_2).

However, $h = x_2 - x_1 = x_1 - x_0$, so

$$(x_2 - x_1)^2 - (x_0 - x_1)^2 = (x_2 - x_1)^4 - (x_0 - x_1)^4 = 0$$

whereas

$$(x_2 - x_1)^3 - (x_0 - x_1)^3 = 2h^3 \quad \text{and} \quad (x_2 - x_1)^5 - (x_0 - x_1)^5 = 2h^5$$

Consequently, Eq. (7.1) can be rewritten as

$$\int_{x_0}^{x_2} f(x)dx = 2hf(x_1) + \frac{h^3}{3}f''(x_1) + \frac{f^{(4)}(\xi_1)}{60}h^5.$$

If we now replace $f''(x_1)$ by the approximation given in Eq. (2.1), we have

$$\int_{x_0}^{x_2} f(x)dx = 2hf(x_1) + \frac{h^3}{3}\left\{\frac{1}{h^2}\left[f(x_0) - 2f(x_1) + f(x_2)\right] - \frac{h^2}{12}f^{(4)}(\xi_2)\right\}$$
$$+\frac{f^{(4)}(\xi_1)}{60}h^5$$
$$= \frac{h}{3}\left[f(x_0) + 4f(x_1) + f(x_2)\right] - \frac{h^5}{12}\left[\frac{1}{3}f^{(4)}(\xi_2) - \frac{1}{5}f^{(4)}(\xi_1)\right]$$

Alternative methods can be used to demonstrate that the values ξ_1 and ξ_2 in this expression can be replaced by a common value ξ in (x_0, x_2). This is where Simpson's rule comes into play.

Simpson's Rule[8]:

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

When applied to any polynomial of degree three or less, Simpson's formula produces accurate results since the error term involves the fourth derivative of f.

8 Example

Compare the Trapezoidal rule and Simpson's rule approximations to $\int_1^3 f(x)dx$ when f(x) is x^3

Solution on [1, 3] the Trapezoidal and Simpson's rule have the forms

$$Trapezoid: \int_1^3 f(x)dx \approx f(1) + f(3)$$

and

$$Simpson's: \int_1^3 f(x)dx \approx \frac{1}{3}[f(1) + 4f(2) + f(3)].$$

When $f(x) = x^3$ they give

Let $\xi = 2$ in (1, 3)

$$\text{Trapezoid: } \int_1^3 f(x)dx \approx 1^3 + 3^3 = 28 \text{ and}$$

$$\text{Simpson's: } \int_1^3 f(x)dx \approx \frac{1}{3}\left[\left(1^3\right) + 4 \cdot 2^3 + 3^3\right] = 20.$$

The approximation from Simpson's rule is exact because its truncation error involves $f^{(4)}$, which is identically 0 when $f(x) = x^3$

Table 2 summarizes the findings for the function in three locations. It's worth noting that Simpson's Rule is far superior.

Table 2:

$f(x)$	x^3
Exact value	20
Trapezoidal	28
Simpson's	20

9 Composite-Rules

Most of the results of this section can be found in [6].

Gautschi refers to these formulas as "the workhorses of numerical integration." When the interval is finite and the integrand is unproblematic, they will accomplish the job. On infinite intervals, the trapezoidal rule can be surprisingly successful.

Both rules are obtained by applying the simplest kind of interpolation on subintervals of the decomposition

$$a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b, \quad x_k = a + kh, \quad h = \frac{b-a}{n}$$

of the interval $[a, b]$. In the trapezoidal rule, one interpolates linearly on each subinterval $[x_k, x_{k+1}]$, and obtains

$$\int_{x_k}^{x_{k+1}} f(x)dx = \int_{x_k}^{x_{k+1}} \left(L_1 f\right)(x)dx + \int_{x_k}^{x_{k+1}} \left(R_1 f\right)(x)dx, \quad f \in C^1[a, b]$$

where

$$\left(L_1 f\right)(x) = f_k + \left(x - x_k\right) f\left[x_k, x_{k+1}\right].$$

Integrating, we have

$$\int_{x_k}^{x_{k+1}} f(x)dx = \frac{h}{2}\left(f_k + f_{k+1}\right) + R_1(f),$$

where (using Peano Theorem)

$$R_1(f) = \int_{x_k}^{x_{k+1}} K_1(t) f''(t)dt,$$

and

$$\begin{aligned} K_1(t) &= \frac{\left(x_{k+1} - t\right)^2}{2} - \frac{h}{2}\left[\left(x_k - t\right)_+ + \left(x_{k+1} - t\right)_+\right] \\ &= \frac{\left(x_k - t\right)^2}{2} - \frac{h\left(x_{k+1} - t\right)}{2} \\ &= \frac{1}{2}\left(x_{k+1} - t\right)\left(x_k - t\right) \leq 0. \end{aligned}$$

So

$$R_1(f) = -\frac{h^3}{12} f''(\xi_k), \quad \xi_k \in (x_k, x_{k+1})$$

and

$$\int_{x_k}^{x_{k+1}} f(x)dx = \frac{h}{2}(f_k + f_{k+1}) - \frac{1}{12} h^3 f''(\xi_k).$$

This is the elementary trapezoidal rule. Summing over all subinterval gives the rule or the composite trapezoidal rule.

$$\int_a^b f(x)dx = h\left(\frac{1}{2}f_0 + f_1 + \cdots + f_{n-1} + \frac{1}{2}f_n\right) - \frac{1}{12}h^3 \sum_{k=0}^{n-1} f''(\xi_k).$$

Since f'' is continuous on $[a, b]$, the remainder term could be written as

$$R_{1,n}(f) = -\frac{(b-a)h^2}{12} f''(\xi) = -\frac{(b-a)^3}{12n^2} f''(\xi).$$

Since f'' is bounded in absolute value on $[a, b]$ we have

$$R_{1,n}(f) = O\left(h^2\right),$$

when $h \to 0$ and so the composite trapezoidal rule converges when $h \to 0$ (or equivalently, $n \to \infty$), provided that $f \in C^2[a, b]$.

Exercise The following integral is given:

$$\int_{1.3}^{4.3} 5xe^{-2x} dx$$

a) Estimate the value of this integral using the composite trapezoidal rule. Three segments should be used.

b) Find the true error E_t for part (a).

$$a)\int_a^b f(x)dx = \frac{b-a}{2n}\left[f(a) + 2\sum_{i=1}^{n-1} f(a+ih) + f(b)\right]$$

$$h = \frac{b-a}{n} = \frac{4.3-1.3}{3} = 1$$

$$\int_{1.3}^{4.3} f(x)dx \simeq \frac{1}{2}\left[f(1.3) + 2\sum_{i=1}^{3-1} f(1.3+i\cdot 1) + f(4.3)\right]$$

$$= \frac{1}{2}\left[f(1.3) + 2\sum_{i=1}^{2} f(1.3+i\cdot 1) + f(4.3)\right]$$

$$= \frac{1}{2}\left[f(1.3) + 2f(1.3+(1)\cdot 1) + 2f(1.3+(2)\cdot 1) + f(4.3)\right]$$

$$= 0.5[f(1.3) + 2f(2.3) + 2f(3.3) + f(4.3)]$$

$$= 0.5\left[5(1.3)e^{-2(1.3)} + 2(5)(2.3)e^{-2(2.3)} + 2(5)(3.3)e^{-2(3.3)} + 5(4.3)e^{-2(4.3)}\right] =$$

$$= 0.5\left[0.4827 + 0.2311 + 0.0448 + 0.0039\right]$$

$$= 0.3812$$

$$b)\int_{1.3}^{4.3} 5xe^{-2x}dx = 0.3320$$

$$E_t = 0.3320 - 0.3812 = -0.0492$$

The composite Simpson's formula is obtained by using quadratic interpolation across two successive intervals instead of linear interpolation. Its "basic" form, known as Simpson's rule or Simpson formula is

$$\int_{x_k}^{x_{k+1}} f(x)dx = \frac{h}{3}\left(f_k + 4f_{k+1} + f_{k+2}\right) - \frac{1}{90}h^5 f^{(h)}\left(\xi_k\right),$$

$$x_k \leq \xi_k \leq x_{k+1}$$

where it has been assumed that $f \in C^4[a,b]$.

Let us prove the formula for the remainder of Simpson rule. Since the degree of exactness is 3 , Peano theorem yields to

$$R_2(f)=\int_{x_k}^{x_{k+2}} K_2(t) f^{(4)}(t) \mathrm{d} t$$

where

$$K_2(t)=\frac{1}{3!}\left\{\frac{\left(x_{k+1}-t\right)^4}{4}-\frac{h}{3}\left[\left(x_k-t\right)_{+}^3+4\left(x_{k+1}-t\right)_{+}^3+\left(x_{k+2}-t\right)_{+}^3\right]\right\},$$

that is,

$$K_2(t)=\frac{1}{6}\begin{cases}\frac{\left(x_{k+2}-t\right)^4}{4}-\frac{h}{3}\left[4\left(x_{k+1}-t\right)^3+\left(x_{k+2}-t\right)^3\right], & t \in\left[x_k, x_{k+1}\right] \\ \frac{\left(x_{k+2}-t\right)^4}{4}-\frac{h}{3}\left(x_{k+2}-t\right)^3, & t \in\left[x_{k+1}, x_{k+2}\right]\end{cases}$$

One easily checks that for $t \in[a, b], K_2(t) \leq 0$, so we can apply Peano's Theorem.

$$\begin{aligned}
R_2(f) &= \frac{1}{4!} f^{(4)}\left(\xi_k\right) R_2\left(e_4\right) \\
R_2\left(e_4\right) &= \frac{x_{k+2}^5-x_k^5}{5}-\frac{h}{3}\left[x_k^4+4 x_{k+1}^4+x_{k+1}^4\right] \\
&= h\left[2 \frac{x_{k+2}^4+x_{k+2}^3 x_k+x_{k+2}^2 x_k^2+x_{k+2} x_k^3+x_k^4}{5}\right. \\
&\left.-\frac{5 x_k^4+4 x_k^3 x_{k+2}+6 x_k^2 x_{k+2}^2+4 x_k x_{k+2}^3+5 x_{k+2}^4}{12}\right] \\
&= \frac{h}{60}\left(-x_k^4+4 x_k^3 x_{k+2}+6 x_k^2 x_{k+2}^2+4 x_k x_{k+2}^3-x_{k+2}^4\right) \\
&= -\frac{h}{60}\left(x_{k+2}-x_k\right)^4=-4 \frac{h^5}{15}
\end{aligned}$$

Thus,

$$R_2(f)=-\frac{h^5}{90} f^{(4)}\left(\xi_k\right)$$

For the composite Simpson [2] rule we get

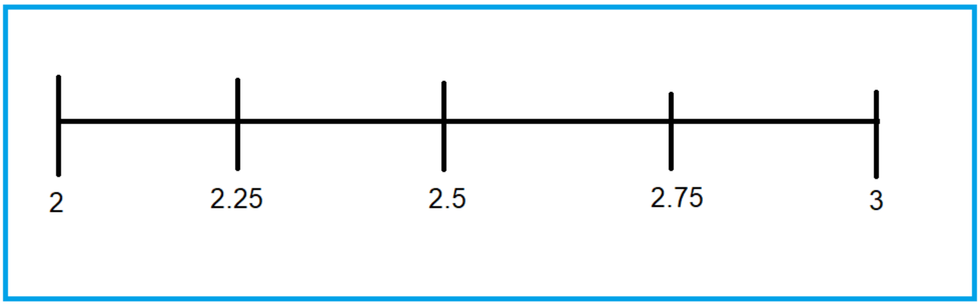
$$\int_a^b f(x) d x=\frac{h}{3}\left(f_0+4 f_1+2 f_2+4 f_3+2 f_4+\cdots+4 f_{n-1}+f_n\right)+R_{2, n}(f)$$

with

$$R_{2,n}(f) = -\frac{1}{180}(b-a)h^4 f^{(4)}(\xi) = -\frac{(b-a)^5}{2880n^4} f^{(4)}(\xi), \quad \xi \in (a,b)$$

One notes that $R_{2,n}(f) = O\left(h^4\right)$, which assures convergence when $n \to \infty$. In the order of accuracy, we also have a gain of one. This is why Simpson's rule has been, and continues to be, one of the most widely used general-purpose integration techniques.

<u>Exercise</u> The integral is as follows: $\int_2^3 x^2 dx$ and n=4. Using composite Simpson's Rule, find the value of the integral.



When n=4 then $h = \frac{3-2}{4}$. The approximation is:

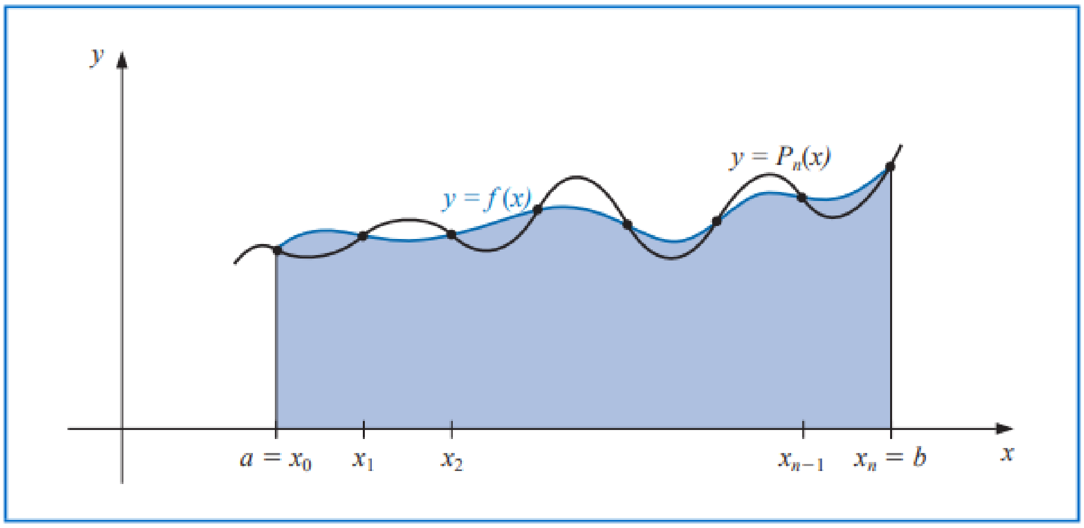
$$\int_2^3 x^2 dx \approx \frac{1/4}{3}\left[y_0 + y_4 + 4\left(y_1 + y_3\right) + 2y_2\right] =$$
$$= \frac{0.25}{3}[f(2) + f(3) + 4\{f(2.75) + f(2.25)\} + 2f(2.5)] =$$
$$= \frac{0.25}{3}[4 + 9 + 4(7.5625 + 5.0625) + 2 \cdot 6.25] =$$
$$= \frac{0.25}{3} \cdot 76 =$$
$$= 6.333$$

10 Closed-Newton-Cotes

Most of the results of this section can be found in [8].

The (n+1) - point closed Newton-Cotes formula uses nodes $x_i = x_0 + ih$, for i = 0, 1, ... , n, where $x_0 = a, x_n = b$ and $h = (b-a)/n$ (See Figure 5). Because the endpoints of the closed interval [a, b] are included as nodes, it is called closed.

Figure 5:



The formula assumes the form

$$\int_a^b f(x)dx \approx \sum_{i=0}^{n} a_i f(x_i) \text{ [8]}$$

where

$$a_i = \int_{x_0}^{x_n} L_i(x)dx = \int_{x_0}^{x_n} \prod_{\substack{j=0 \\ j\neq i}}^{n} \frac{(x-x_j)}{(x_i-x_j)}dx$$

When n=1 we have Trapezoidal rule, when n=2 we have Simpson's rule, when n=3 we have Simpson's Three-Eighths rule.

n=1: Trapezoidal Rule

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

Exercise Aproximate the integral $\int_0^{\pi/4} \cos x dx$ using Closed Newton Cotes with n=1.

$$\int_0^{\pi/4} \cos x dx \approx 0.70710$$

$$\int_0^{\pi/4} \cos x dx = \frac{h}{2}\left[f\left(x_0\right) + f\left(x_1\right)\right]$$

$$= \frac{\pi/4}{2}[\cos(0) + \cos(\pi/4)]$$

$$\approx 0.67037$$

11 Open-Newton-Cotes

Most of the results of this section can be found in [8].

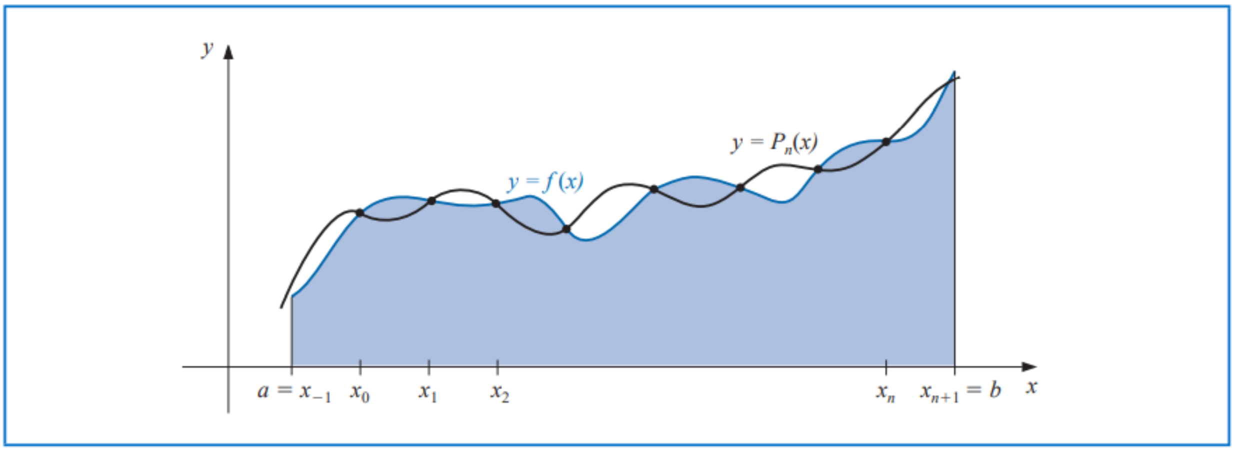
The endpoints of [a, b] are not included as nodes in the open Newton-Cotes formulae. They use the nodes $x_i = x_0 + ih$, for each i = 0, 1, ... , n, where $h = (b - a)/(n + 2)$ and $x_0 = a + h$. This implies that $x_n = b - h$, as a result, the endpoints are labeled by setting $x_{-1} = a$ and $x_{n+1} = b$, as shown in Figure 6. Open formulas contain all the nodes used for the approximation within the open interval (a, b). The formulas become

$$\int_a^b f(x)dx = \int_{x_{-1}}^{x_{n+1}} f(x)dx \approx \sum_{i=0}^{n} a_i f\left(x_i\right) [8]$$

where

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

Figure 6:



When n=0 we have Midpoint rule.

n=0: Midpoint Rule

$$\int_{x_{-1}}^{x_1} f(x)dx = 2hf\left(x_0\right) + \frac{h^3}{3} f''(\xi), \quad \text{where} \quad x_{-1} < \xi < x_1$$

<u>**Exercise**</u> Aproximate the integral $\int_0^{\pi/3} \cos x dx$ using Open Newton Cotes with n=0.

$$\int_0^{\pi/3} \cos x dx \approx 0.866025$$

$$I = 2hf\left(x_0\right)$$

$$h = \frac{b-a}{n+2} = \frac{\pi/3 - 0}{2} = \pi/6$$

$$x_0 = a + h = 0 + \pi/6$$

$$x_0 = \pi/6$$

$$I = 2(\pi/6)\cos(\pi/6) \approx 0.90689$$

12 Romberg Integration

Most of the results of this section can be found in [7] and [8].

We will show how Richardson extrapolation applied to Composite Trapezoidal rule outcomes may be used to achieve high accuracy approximations with minimum processing expense in this section.

In Section 9 we discovered that the Composite Trapezoidal rule has a truncation error of order $O(h^2)$. Specifically, we demonstrated that for $h = (b-a)/n$ and $x_j = a + jh$ we obtain

$$\int_a^b f(x)dx = \frac{h}{2}\left[f(a) + 2\sum_{j=1}^{n-1} f\left(x_j\right) + f(b)\right] - \frac{(b-a)f''(\mu)}{12}h^2$$

for some number μ in (a, b). By an alternative method it can be shown, that if $f \in C^\infty[a,b]$, the Composite Trapezoidal rule can also be written with an error term in the form

$$\int_a^b f(x)dx = \frac{h}{2}\left[f(a) + 2\sum_{j=1}^{n-1} f\left(x_j\right) + f(b)\right] + K_1h^2 + K_2h^4 + K_3h^6 + \cdots,$$

where each K_i is a constant that depends only on $f^{(2i-1)}(a)$ and $f^{(2i-1)}(b)$.

Remember from Section 3 that Richardson extrapolation can be used on any approximation procedure with a truncation error of the form

$$\sum_{j=1}^{m-1} K_j h^{\alpha_j} + O\left(h^{\alpha_m}\right),$$

for a collection of constants K_j and when $\alpha_1 < \alpha_2 < \alpha_3 < \cdots < \alpha_m$. We demonstrated the effectiveness of these strategies in that section when the approximation procedure has a truncation error with only even powers of h, i.e. when the truncation error has the form.

$$\sum_{j=1}^{m-1} K_j h^{2j} + O\left(h^{2m}\right).$$

The Composite Trapezoidal rule is an obvious candidate for extrapolation because it has this form. As a result, a technique known as Romberg integration is created.

We use the results of the Composite Trapezoidal rule with n = 1, 2, 4, 8, 16, ... to approximate the integral $\int_a^b f(x)dx$, and denote the resulting approximations, respectively, by $R_{1,1}, R_{2,1}, R_{3,1}$, etc. Extrapolation is then used in the way described in Section 3, that is, we obtain $O\left(h^4\right)$ approximations $R_{2,2}, R_{3,2}, R_{4,2}$, etc., by

$$R_{k,2} = R_{k,1} + \frac{1}{3}\left(R_{k,1} - R_{k-1,1}\right), \quad \text{for } k = 2, 3, \ldots$$

Then $O\left(h^6\right)$ approximations $R_{3,3}, R_{4,3}, R_{5,3}$, etc., by

$$R_{k,3} = R_{k,2} + \frac{1}{15}\left(R_{k,2} - R_{k-1,2}\right), \quad \text{for } k = 3, 4, \ldots$$

In general, after obtaining the required appropriate approximations from the $R_{k,j-1}$, we determine the $O\left(h^{2j}\right)$ approximations from

$$R_{k,j} = R_{k,j-1} + \frac{1}{4^{j-1}-1}\left(R_{k,j-1} - R_{k-1,j-1}\right), \quad \text{for } k = j, j+1, \ldots [8]$$

MATLAB Example[7]

function [I,nfev]=Romberg(f,a,b,epsi,nmax)

%ROMBERG - approximate an integral using Romberg method

%call [I,NFEV]=ROMBERG(F,A,B,EPSI,NMAX)

%F - integrands

%A,B - integration limits

%EPSI - tolerance

```
%NMAX - maximum number of iterations
%I - approximate integral value
%NFEV - number of function evaluations


if nargin < 5
   nmax=10;
end
if nargin < 4
   epsi=1e-3;
end
R=zeros(nmax,nmax);
h=b-a;
% first iteration
R(1,1)=h/2*(sum(f([a,b])));
nfev=2;
for k=2:nmax
   %trapezes formula
   x=a+((1:2^(k-2))-0.5)*h;
   R(k,1)=0.5*(R(k-1,1)+h*sum(f(x)));
   nfev=nfev+length(x);
   %extrapolation
   plj=4;
   for j=2:k
      R(k,j)=(plj*R(k,j-1)-R(k-1,j-1))/(plj-1);
      plj=plj*4;
```

```
    end
    if (abs(R(k,k)-R(k-1,k-1))<epsi)&&(k>3) %success
        I=R(k,k);
        %R(1:k,1:k)
        return
    end
    %halving step
    h=h/2;
end
error('iteration number exceeded')
```

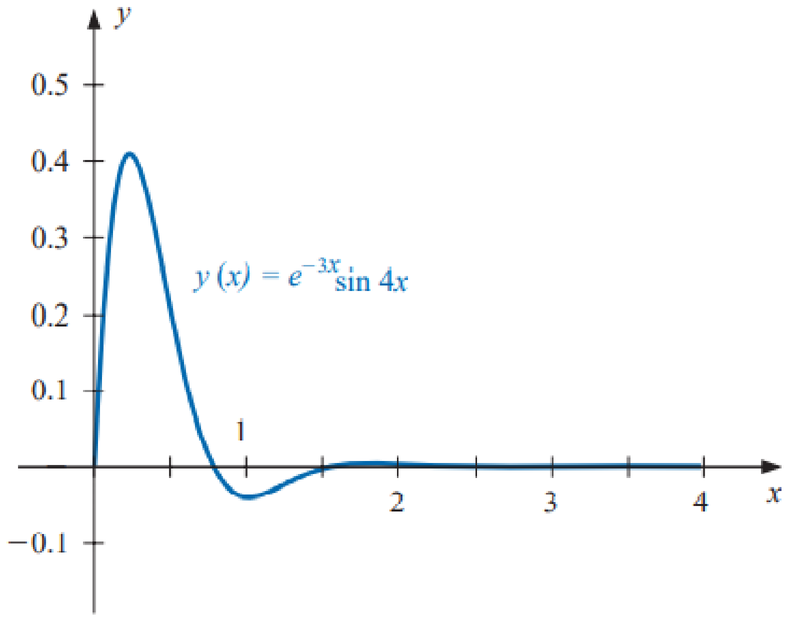
13 Adaptive Quadrature

Most of the results of this section can be found in [8].

In most cases, composite formulas are quite successful, however they occasionally suffer from the requirement of equally-spaced nodes. This is inefficient when integrating a function on an interval that contains both regions with large functional variation and regions with minimal functional variation.

$y(x) = e^{-3x} \sin 4x$ is the unique solution to the differential equation $y'' + 6y' + 25 = 0$ that additionally satisfies y(0) = 0 and $y'(0) = 4$. This sort of function is common in mechanical engineering because it describes certain characteristics of spring and shock absorber systems, and it is also popular in electrical engineering since it is a common solution to basic circuit difficulties. Figure 7 illustrates the graph of y(x) for x in the range [0, 4].

Figure 7:



Let's say we need the integral of y(x) on the range [0, 4]. The graph suggests that the integral on [3, 4] must be extremely near to 0, and that the integral on [2, 3] should similarly be small. However, the function varies significantly between [0, 2], therefore it is unclear what the integral is on this interval. This is an example of why composite integration isn't the best solution. On [2, 4], a very low-order method could be employed, but on [0, 2], a higher-order method would be required.

In this section, we'll look at the following question:

- How do we know which technique to use on different parts of the integration interval, and how precise can the final approximation be?

We will see that we can both answer this question and determine approximations that satisfy certain accuracy requirements under fairly realistic conditions.

If an integral's approximation error on a certain interval is to be uniformly distributed, the large-variation regions require a smaller step size than the low-variation regions. For this type of problem, an effective approach should forecast the amount of functional variation and adjust the step size as needed. Adaptive quadrature methods are the name for these techniques. Adaptive methods are very common in professional software packages because, in addition to being efficient, they typically yield approximations that are within a specific tolerance.

In this part, we'll look at how an adaptive quadrature method can be used to reduce approximation error and anticipate an error estimate for the approximation that doesn't require knowledge of the function's higher derivatives. The strategy we'll go over is based on the Composite Simpson's rule, but it can easily be adjusted to work with different composite techniques.

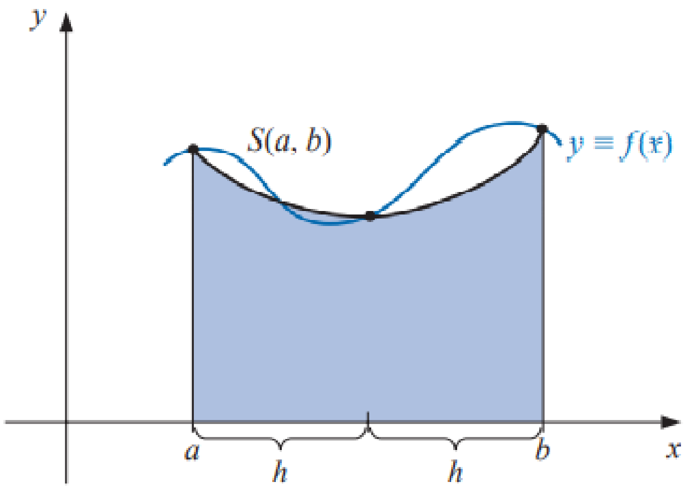
Assume that we need to approximate $\int_a^b f(x)dx$ to within a certain tolerance $\epsilon > 0$. The first step is to use Simpson's rule with step size h = (b - a)/2. This results in (see Figure 8)

$$\int_a^b f(x)dx = S(a,b) - \frac{h^5}{90}f^{(4)}(\xi), \quad \text{for some } \xi \text{ in } (a,b) \tag{13.1}$$

where we the Simpson's rule approximation on [a, b] is denoted by

$$S(a,b) = \frac{h}{3}[f(a) + 4f(a+h) + f(b)]$$

Figure 8:



The next step is to find an approximation for accuracy that does not require $f^{(4)}(\xi)$. To accomplish this, we use the Composite Simpson's rule with n $=$ 4 and step size $(b-a)/4 = h/2$, giving

$$\int_a^b f(x)dx = \frac{h}{6}\left[f(a) + 4f\left(a+\frac{h}{2}\right) + 2f(a+h) + 4f\left(a+\frac{3h}{2}\right) + f(b)\right] - \left(\frac{h}{2}\right)^4 \frac{(b-a)}{180} f^{(4)}(\tilde{\xi}) \qquad (13.2)$$

for some $\tilde{\xi}$ in (a, b). To simplify notation, let

$$S\left(a, \frac{a+b}{2}\right) = \frac{h}{6}\left[f(a) + 4f\left(a+\frac{h}{2}\right) + f(a+h)\right]$$

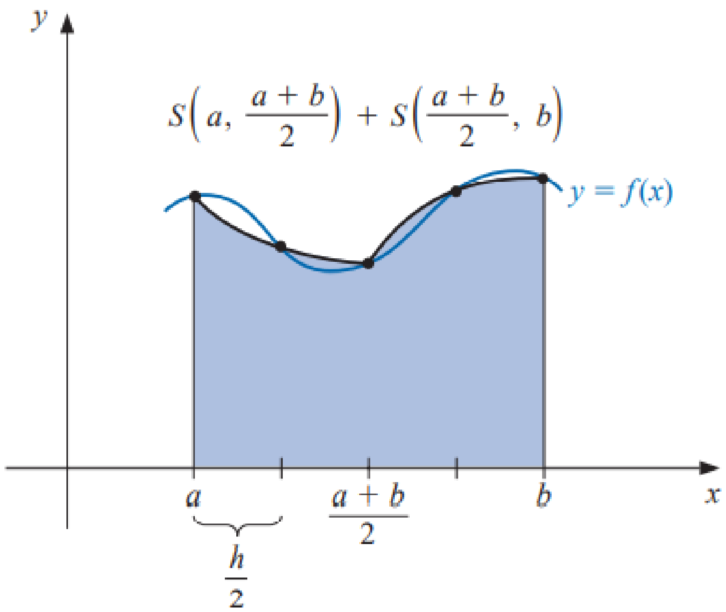
and

$$S\left(\frac{a+b}{2}, b\right) = \frac{h}{6}\left[f(a+h) + 4f\left(a+\frac{3h}{2}\right) + f(b)\right]$$

Then Eq.(13.2) can be rewritten (see Figure 9) as

$$\int_a^b f(x)dx = S\left(a, \frac{a+b}{2}\right) + S\left(\frac{a+b}{2}, b\right) - \frac{1}{16}\left(\frac{h^5}{90}\right) f^{(4)}(\tilde{\xi}). \quad (13.3)$$

Figure 9:



The error estimation is based on the premise that $\xi \approx \tilde{\xi}$ or, more accurately, that $f^{(4)}(\xi) \approx f^{(4)}(\tilde{\xi})$, and the technique's effectiveness is dependent on the accuracy of this assumption. If it is accurate, then equating the integrals in Eqs. (13.1) and (13.3) gives

$$S\left(a, \frac{a+b}{2}\right) + S\left(\frac{a+b}{2}, b\right) - \frac{1}{16}\left(\frac{h^5}{90}\right) f^{(4)}(\xi) \approx S(a,b) - \frac{h^5}{90} f^{(4)}(\xi)$$

so

$$\frac{h^5}{90} f^{(4)}(\xi) \approx \frac{16}{15}\left[S(a,b) - S\left(a, \frac{a+b}{2}\right) - S\left(\frac{a+b}{2}, b\right)\right]$$

Using this estimate in Eq. (13.3) produces the error estimation

$$\left|\int_a^b f(x)dx - S\left(a, \frac{a+b}{2}\right) - S\left(\frac{a+b}{2}, b\right)\right| \approx \frac{1}{16}\left(\frac{h^5}{90}\right) f^{(4)}(\xi)$$

$$\approx \frac{1}{15}\left|S(a,b) - S\left(a, \frac{a \mid b}{2}\right) - S\left(\frac{a \mid b}{2}, b\right)\right|$$

This implies that $S(a, (a+b)/2) + S((a+b)/2, b)$ approximates $\int_a^b f(x)dx$ about 15 times better than it agrees with the computed value S(a, b). Thus, if

$$\left|S(a,b) - S\left(a, \frac{a+b}{2}\right) - S\left(\frac{a+b}{2}, b\right)\right| < 15\varepsilon \quad [8]$$

we expect to have

$$\left|\int_a^b f(x)dx - S\left(a, \frac{a+b}{2}\right) - S\left(\frac{a+b}{2}, b\right)\right| < \varepsilon \quad [8]$$

and

$$S\left(a, \frac{a+b}{2}\right) + S\left(\frac{a+b}{2}, b\right) \quad [8]$$

is expected to be a sufficiently accurate approximation to $\int_a^b f(x)dx$.

Exercise Examine the error estimate's accuracy when applied to the integral:

$$\int_0^{\pi/4} \cos x dx = \frac{\sqrt{2}}{2}$$

by comparing

$$\frac{1}{15}\left|S\left(0,\frac{\pi}{4}\right) - S\left(0,\frac{\pi}{8}\right) - S\left(\frac{\pi}{8},\frac{\pi}{4}\right)\right| \quad \text{to} \quad \left|\int_0^{\pi/4} \cos x dx - S\left(0,\frac{\pi}{8}\right) - S\left(\frac{\pi}{8},\frac{\pi}{4}\right)\right|.$$

We have

$$S\left(0,\frac{\pi}{4}\right) = \frac{\pi/8}{3}\left[\cos 0 + 4\cos\frac{\pi}{8} + \cos\frac{\pi}{4}\right] = \frac{\pi}{24}\cdot 5.4026249 = 0.707201$$

and

$$S\left(0,\frac{\pi}{8}\right)+S\left(\frac{\pi}{8},\frac{\pi}{4}\right) = \frac{\pi/16}{3}\left[\cos 0 + 4\cos\frac{\pi}{16} + 2\cos\frac{\pi}{8} + 4\cos\frac{3\pi}{16} + \cos\frac{\pi}{4}\right]$$
$$= 0.707112.$$

So

$$\left|S\left(0,\frac{\pi}{4}\right) - S\left(0,\frac{\pi}{8}\right) - S\left(\frac{\pi}{8},\frac{\pi}{4}\right)\right| = |0.707201 - 0.707112| = 0.002145293$$

The estimate for the error obtained when using S(a,(a+b))+S((a+b), b) to approximate $\int_a^b f(x)$ is consequently

$$\frac{1}{15}\left|S\left(0,\frac{\pi}{2}\right) - S\left(0,\frac{\pi}{4}\right) - S\left(\frac{\pi}{4},\frac{\pi}{2}\right)\right| = 0.000089$$

which closely approximates the actual error

$$\left|\int_0^{\pi/4} \cos x dx - 0.707112\right| = 0.00000521$$

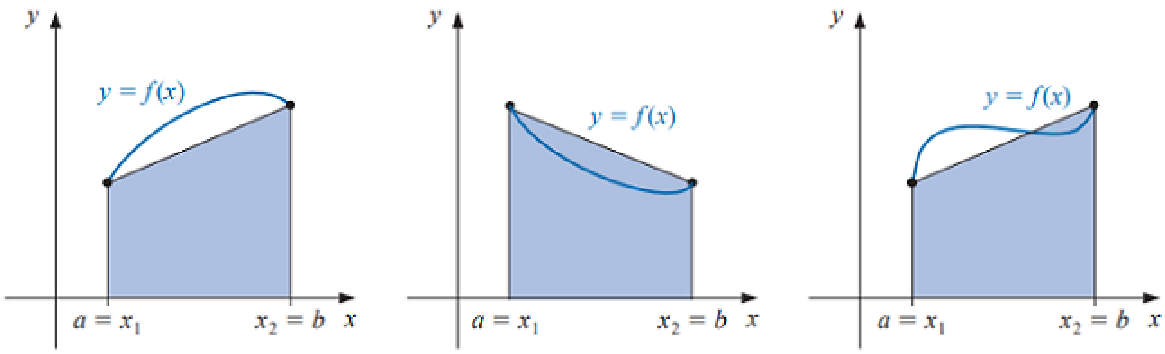
14 Gaussian Quadrature

Most of the results of this section can be found in [8].

In Chapters 10 and 11, the Newton-Cotes formulae were obtained by integrating interpolating polynomials. Because the (n + 1)st derivative of the function being approximated is involved in the error term in the interpolating polynomial of degree n, the Newton-Cotes formula is precise when approximating the integral of any polynomial of degree less than or equal to n.

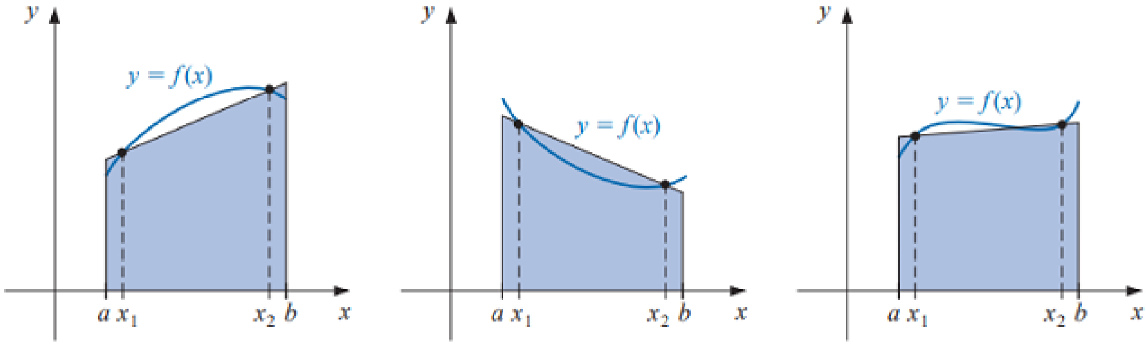
The Newton-Cotes formulas all use function values at evenly spaced points. This restriction is useful when the formulae are combined to generate the composite rules we discussed in Chapter 9, but it can reduce the approximation's accuracy greatly. Consider the Trapezoidal rule, which was used to find the integrals of the functions whose graphs are given in Figure 10.

Figure 10:



By integrating the linear function that connects the endpoints of the function's graph, the Trapezoidal rule approximates the integral. However, this is unlikely to be the most accurate line for estimating the integral. In the vast majority of cases, lines like those shown in Figure 11 would provide considerably better estimates.

Figure 11:



Gaussian quadrature selects points for evaluation that are optimally spaced rather than evenly spaced. The coefficients $c_1, c_2, \ldots, c_n$ and the nodes $x_1, x_2, \ldots, x_n$ in the interval [a, b] , are chosen to minimize the expected error produced in the approximation

$$\int_a^h f(x)dx \approx \sum_{i=1}^{n} c_i f\left(x_i\right)$$

To calculate this precision, we assume that the best combination of these values yields the precise solution for the largest class of polynomials, i.e., the choice with the highest precision.

The approximation formula's coefficients $c_1, c_2, \ldots, c_n$ are arbitrary, while the nodes $x_1, x_2, \ldots, x_n$ are only constrained by the fact that they must exist in [a, b], the integration interval. This gives us a total of 2n parameters from which to choose. If a polynomial's coefficients are considered parameters, the class of polynomials with a degree of at most 2n - 1 also contains 2n parameters. This is the largest class of polynomials for which a precise formula is realistic to predict. Exactness on this set can be achieved by selecting the right values and constants.

We'll show how to choose the coefficients and nodes for n = 2 and the integration interval is [-1, 1] to illustrate the technique for selecting the suitable parameters. The technique will then be adjusted when integrating over an arbitrary interval for an arbitrary choice of nodes and coefficients.

Suppose we want to determine c_1, c_2, x_1 , and x_2 so that the integration formula

$$\int_{-1}^{1} f(x)dx \approx c_1 f(x_1) + c_2 f(x_2)$$

gives the exact result whenever f (x) is a polynomial of degree 2(2) - 1 = 3 or less, that is, when

$$f(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3,$$

for some collection of constants, $a_0, a_1, a_2,$ and a_3 . Because

$$\int \left(a_0 + a_1 x + a_2 x^2 + a_3 x^3\right) dx = a_0 \int 1dx + a_1 \int xdx + a_2 \int x^2 dx + a_3 \int x^3 dx$$

this is equivalent to showing that the formula gives exact results when f (x) is $1, x, x^2$ and x^3 . Hence, we need c_1, c_2, x_1 and x_2, so that

$$c_1 \cdot 1 + c_2 \cdot 1 = \int_{-1}^{1} 1dx = 2, \qquad c_1 \cdot x_1 + c_2 \cdot x_2 = \int_{-1}^{1} xdx = 0,$$
$$c_1 \cdot x_1^2 + c_2 \cdot x_2^2 = \int_{-1}^{1} x^2 dx = \tfrac{2}{3}, \quad \text{and} \quad c_1 \cdot x_1^3 + c_2 \cdot x_2^3 = \int_{-1}^{1} x^3 dx = 0.$$

A little algebra shows that this system of equations has the unique solution

$$c_1 = 1, \quad c_2 = 1, \quad x_1 = -\frac{\sqrt{3}}{3}, \quad \text{and} \quad x_2 = \frac{\sqrt{3}}{3}$$

which gives the approximation formula

$$\int_{-1}^{1} f(x)dx \approx f\left(\frac{-\sqrt{3}}{3}\right) + f\left(\frac{\sqrt{3}}{3}\right)$$

This formula has degree of precision 3, which means it gives the precise answer for any polynomial of degree 3 or less.

15 Numerical Integration in MATLAB

Most of the results of this section can be found in [6].

Quad and quadl are the two main functions for numerical integration in MATLAB. Both require finite values for a and b, as well as the integrand having no singularities on [a, b]. Change of variable, integration by parts, and analytic treatment of the integral over part of the range are some of the ways that can be used to construct an integral that can be handled by quad and quadl for infinite integrals and integrals with singularities.

q = quad(fun,a,b,tol) (and similarly for quadl) is the most common usage, where fun denotes the function to be integrated. A string, an inline object, or a function handle can all be used. The tol argument specifies the absolute error tolerance, which is set to 1e-6 by default. A small multiple of eps times an estimate of the integral is a suggested number. The values of [fcount a b-a Q] calculated during the recursion are shown in the form q = quad(fun,a,b,tol,trace) with a nonzero trace.

[q,fcount] = quad(...) returns the number of function evaluations.

Suppose we want to approximate $\int_0^{\pi} x \sin x \, \mathrm{d}x$. We can store the integrand in an M-file, say xsin.m:

The approximate value is computed by:

» quad(@xsin,0,pi)

ans = 3.1416

As described in Section 13, the quad function is an implementation of a Simpson-type quadrature. quadl is more accurate, with a 4 point Gauss-Lobatto formula (degree of exactness 5) and a 7 point Kronrod extension

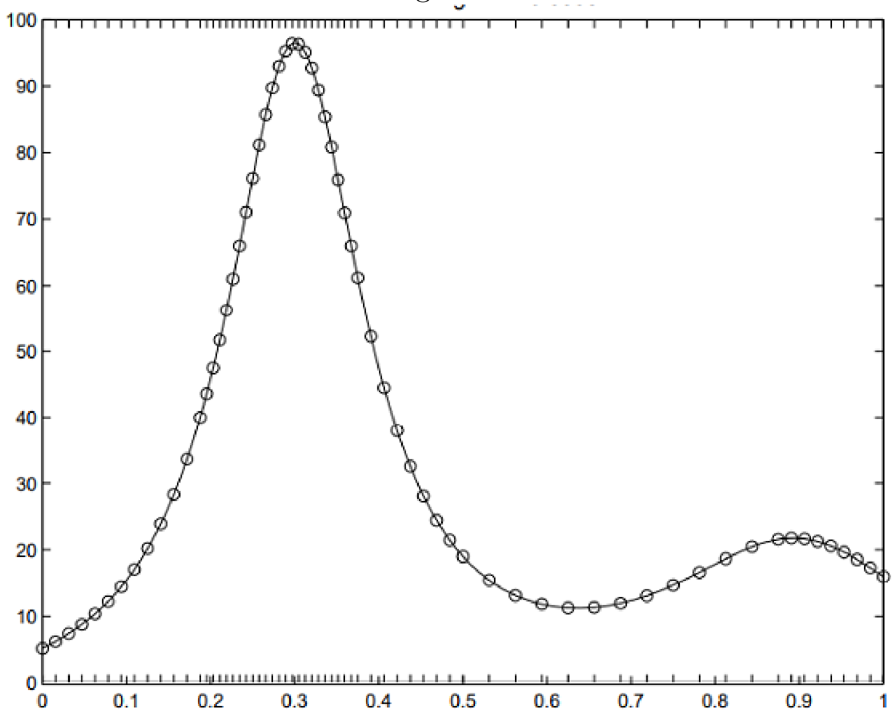
(with degree of exactness 9). Adaptive quadrature is used in both routines. They divide the integration range into subintervals and apply the basic integration rule to each one. They choose the
subintervals according to the local behavior of the integrand, placing the smallest ones where the integrand is changing most rapidly. If the subintervals become too small or if an excessive number of function evaluations are performed, warning messages are generated, indicating that the integrand has a singularity.

To illustrate how quad and quadl we shall approximate

$$\int_0^1 \left(\frac{1}{(x-0.3)^2+0.01} + \frac{1}{(x-0.09)^2+0.04} - 6 \right) dx$$

The integrand is the MATLAB function humps, which are used in MATLAB demos or to test numerical integration methods. With tol=1e-4, we'll apply quad to this function. The integrand is plotted in Figure 12 with tick marks on the x-axis indicating where the integrand was calculated and circles indicating the integrand's corresponding values. The subintervals are smaller where the integrand is rapidly variable, as shown in the graph. We got it by changing the quad MATLAB function.

Figure 12:

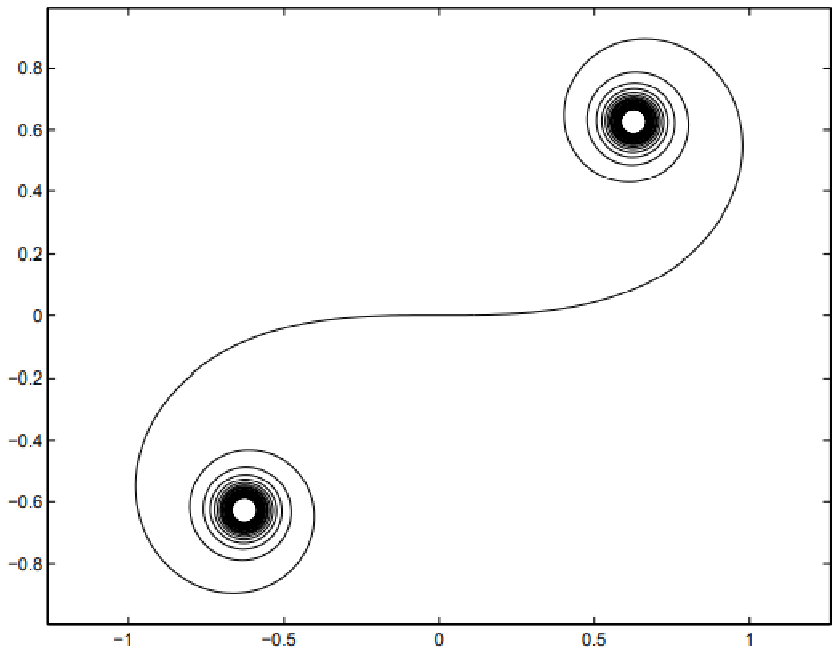


The next example approximates the Fresnel integrals

$$x(t) = \int_0^t \cos\left(u^2\right) \mathrm{d}u, \quad y(t) = \int_0^t \sin\left(u^2\right) \mathrm{d}u$$

The Fresnel spiral is a set of parametric equations for a curve. Its graphical representation is shown in Figure 13, which was created by sampling 1000 equally spaced points on the interval [-4π, 4π]. We take advantage of the symmetry to avoid integrating on [0, t] repeatedly by integrating over each subinterval and then calculating the cumulative sums with cumsum.

Figure 13:



The trapz function is used to integrate functions that are provided by values rather than their analytical expression. It implements the composite trapezoidal rule. The nodes do not have to be evenly spaced. It works well when integrating periodical functions on intervals whose lengths are an integer multiple of the period, as we have seen in Section 6.3.1.

Example: [6]

» x=linspace(0,2*pi,10);

» y=1./(2+sin(x));

» trapz(x,y)

ans = 3.62759872810065

» 2*pi*sqrt(3)/3-ans

ans = 3.677835813675756e-010

The exact value of the integral is $\frac{2}{3}\sqrt{3}\pi$, so the error is less than 10^{-9}.

The function quadv accepts a vector argument and returns a vector.

Quadgk is a more recent function. It uses a Gauss-Kronrod pair to accomplish adaptive quadrature (15th and 7th order formulas). Aside from the approximate value of the integral, it can also return an error bound and take a number of options for controlling the integration process (for example, we can define singularities). It integrates on a straight line within the complex plane if the integrand is complex-valued or the limits are complex. We'll use two as examples. The first computes

$$\int_{-1}^{1} \frac{\sin x}{x} \, \mathrm{d}x$$

» format long

» ff=@(x) sin(x)./x;

» quadgk(ff,-1,1,'RelTol',1e-8,'AbsTol',1e-12)

ans = 1.892166140734366

Notice that quad and quadl fail; they return NaN. Nevertheless, they succeed if we compute the integral as:

$$\int_{-1}^{1} \frac{\sin x}{x} \, \mathrm{d}x = \int_{-1}^{0} \frac{\sin x}{x} \, \mathrm{d}x + \int_{0}^{1} \frac{\sin x}{x} \, \mathrm{d}x$$

The second example uses Waypoints to integrate around a pole using a piecewise linear contour:

» Q = quadgk(@(z)1./(2*z - 1),-1-i,-1-i,'Waypoints',[1-i,1+i,-1+i])

Q =   0.0000 + 3.1416i

16 Aplications

Most of the results of this section can be found in [6].

Computation of an ellipsoid surface

Consider an ellipsoid obtained by rotating the ellipse in Figure 14 around the x axis. The radius ρ is described as a function of axial coordinate by the equation

$$\rho^2(x) = \alpha^2 \left(1 - \beta^2 x^2\right), \quad -\frac{1}{\beta} \leq x \leq \frac{1}{\beta}$$

where α and β are such that $\alpha^2\beta^2 < 1$

For test we chose the following values for the parameters: $\alpha = (\sqrt{2} - 1)/10, \beta = 10$

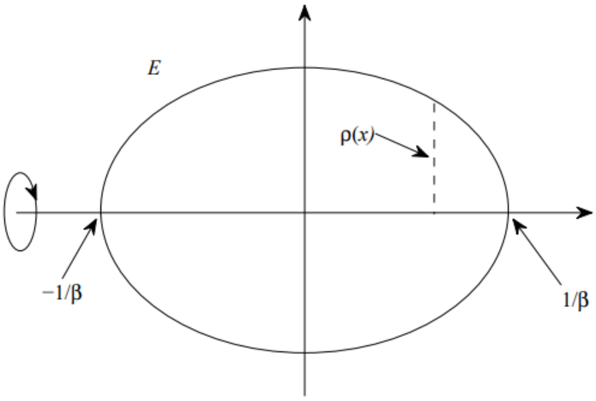
The surface is given by

$$I(f) = 4\pi\alpha \int_0^{1/\beta} \sqrt{1 - K^2x^2}\, \mathrm{d}x$$

where $K^2 = \beta^2\sqrt{1 - \alpha^2\beta^2}$. Since $f'(1/\beta) = -100$, an adaptive quadrature seems to be appropriate.

We can compute the exact value and its floating point approximation using Symbolic Math Toolbox:

Figure 14:



```
clear
syms alpha beta K2 s2 x f vI
s2=sqrt(sym(2));
alpha=(s2-1)/10;
beta=sym(10);
K2=beta^2*sqrt(1-alpha^2*beta^2);
f=sqrt(1-K2*x^2);
vI=4*sym(pi)*alpha*int(f,0,1/beta)
vpa(vI,16)
```

The results are:

```
vI =
1/2 1/2 1/2 1/2 1/2
1/100 pi (-2 (-(-2 + 2 2 ) + 1) + 2 (-(-2 + 2 2 )
1/2 1/2 1/2 3/4 1/2 1/4
+ 1) 2 + (-2 + 2 2 ) asin((-2 + 2 2 ) ))
ans =
0.04234752094082434
```

The next script approximates the surface with a tolerance of 1e-8 using the following functions: Romberg, adquad, and MATLAB quad and quadl.

```
err=1e-8;
beta=10;
alpha=(sqrt(2)-1)/10;
alpha2=alpha^2;
beta2=beta^2;
K2=beta2*sqrt(1-alpha2*beta2);
```

f=@(x) sqrt(1-K2*x.^2);

fpa=4*pi*alpha;

[vi(1), nfe (1)] = Romberg ($f, 0, 1/$ beta, err, 100);

[vi(2), nfe(2)] = adquad ($f, 0, 1/$ beta, err);

[vi(3), nfe (3)] = quad (f,0,1/beta, err)

[vi(4), nfe(4)] = $quadl(f, 0, 1/$ beta, err)

vi=fpa*vi;

meth='Romberg','adquad','quad','quadl';

for i=1:4

fprintf('%8s %18.16f %3d\n',methi,vi(i),nfe(i))

end

Here is the output:

Romberg 0.0423475209214685 129

adquad 0.0423475209189811 65

quad 0.0423475203088494 37

quadl 0.0423475209279265 48

Romberg method is inferior to adaptive quadratures. Surprisingly, quad beats quadl.

Computation of the wind action on a sailboat mast

The sailboat displayed in Fig 15(a) is exposed to the effects of wind force.

The mast is represented by the straight line AB, which is of length L, while BO represents one of the two shrouds (strings for the side stiffening of the mast).

Any infinitesimal element of the sail transfers a force of magnitude $f(x)\mathrm{d}x$

to the corresponding element of length dx of the mast. The following law expresses the change in f together with the height x measured from point A (mast base).

$$f(x) = \frac{\alpha x}{x + \beta} e^{-\gamma x}$$

where α, β and γ are given constants.

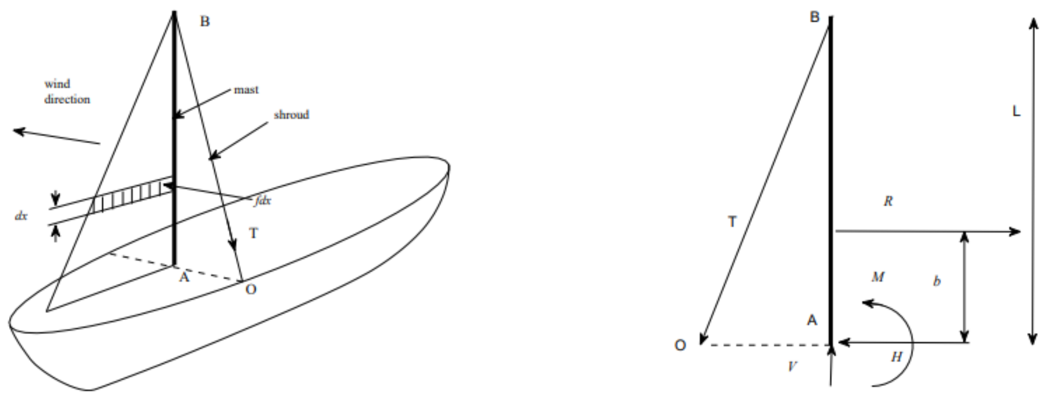
The resultant R of the force f is defined as

$$R = I(f) \equiv \int_0^L f(x) \mathrm{d}x$$

and is applied at distance equal to b (to be determined) from the basis of the mast. The formula for b is $b = I(xf)/I(f)$.

Computing R and b is crucial for the structural design of the mast and shroud section. It is possible to study the hyperstatic structure mastshroud once the values of R and b are known. This analysis has as results the reactions V , H, and M at the basis of the mast and the traction T that is transmitted by the shroud (see Figure 15(b)). The internal activities in the structure, as well as the maximum stresses arising in the mast AB and the shroud BO, can be identified in a subsequent phase, from which, given that the safety verifications are met, the geometric parameters of the sections AB and BO can finally be designed.

Figure 15:



Schematic representation of a sailboat (left); forces acting on the mast (right)

We approximate R and b by adquad and MATLAB quad and quadl functions. The function sailboat in MATLAB Figure 16 computes the approximations and plot number of function evaluations versus minus decimal logarithm of error. We test the function for $\alpha = 50, \beta = 5/3, \gamma = 1/4,$ and $L = 10$. Required tolerances are in the range 10^{-1} to 10^{-9}. The calling command and the results are given below:

$$>> \text{ sailboat } (50, 5/3, 1/4, 10)$$

$$\text{R1} = 100.061368317961$$

$$\text{R2} = 100.061368317941$$

$$\text{R3} = 100.061368317962$$

$$\text{b1} = 4.03145652950332$$

$$\text{b2} = 4.03145652950425$$

$$\text{b3} = 4.03145652950326$$

Figure 16:

MATLAB Source 6.14 Computation of the wind action on a sailboat mast

```
function sailboat(alpha, beta, gamma, L)
%SAILBOAT - computation of wind action on
%           a sailboat mast
% Alfio Quarteroni, Riccardo Sacco, Fausto Saleri
% Numerical  Mathematics
% Springer 2000

f = @(x) alpha*x./(x+beta).*exp(-gamma*x);
xf = @(x) x.*f(x);
x=1:9;
err=10.^(-x);
for k=1:9
    [R1,ne1(k)]= adquad(f,0,L,err(k));
    [b1,neb1(k)]= adquad(xf,0,L,err(k)); b1=b1/R1;
    [R2,ne2(k)]= quad(f,0,L,err(k));
    [b2,neb2(k)]= quad(xf,0,L,err(k)); b2=b2/R2;
    [R3,ne3(k)]=quadl(f,0,L,err(k));
    [b3,neb3(k)]=quadl(xf,0,L,err(k)); b3=b3/R3;
end
subplot(1,2,1)
plot(x,ne1,'b-x',x,ne2,'r-+',x,ne3,'g--d')
xlabel('-log_{10}(err)','FontSize',14); ylabel('n','FontSize',14)
legend('adquad','quad','quadl',0)
title('Computation of \it{R}','FontSize',14)
subplot(1,2,2)
plot(x,neb1,'b-x',x,neb2,'r-+',x,neb3,'g--d')
xlabel('-log_{10}(err)','FontSize',14); ylabel('n','FontSize',14)
legend('adquad','quad','quadl',0)
title('Computation of \it{b}','FontSize',14)
R1,R2,R3
b1,b2,b3
```

References

[1] C.H. Edwards - The Historical Development of the Calculus (p. 99)-Springer-Verlag

[2] Daffodil International University: Application of numerical integration and differentiation in real life, https://www.slideshare.net/ashaf15-7473/application-of-numerical-integration-and-differentiation-in-real-life

[3] Felix Klein - Elementary Mathematics from an Advanced Standpoint. Arithmetic, Algebra, Analysis (p. 209) - Dover Publications

[4] Gheorghe Coman, Ioana Chiorean, Teodora Cătinaş: Numerical Analysis. An Advanced Course.

[5] Petronas Twin Towers, https://www.petronastwintowers.com.my/design-and-structures/

[6] Radu T. Trîmbiţaş: Numerical Analysis in MATLAB, Cluj-Napoca, July 2008

[7] Radu T. Trîmbiţaş: Romberg.m electronic file

[8] Richard L. Burden, J. Douglas Faires: Numerical Analysis, 2010

[9] The Volume of a wine barrel, http://www.matematicasvisuales.com/english/html/history/kepler/keplerbarrel.html

[10] https://www.youtube.com/watch?v=Dx7fDb6boRA&ab_channel=numericalmethodsguy

[11] https://www.chegg.com/homework-help/use-composite-simpson-s-rule-approximate-integrals-exercise-chapter-4.4-problem-3e-solution-9781305253667-exc

[12] https://www.youtube.com/watch?v=ts-zBjsqiX0&ab_channel=DrSajjadKhanMathAcademy