CLASS NOTES
Models, Algorithms and Data: Introduction to computing
2019

Petros Koumoutsakos, Jens Honore Walther, Julija Zavadlav

(Last update: April 25, 2019)

IMPORTANT DISCLAIMERS

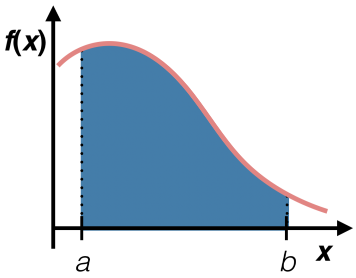
Much of the material (ideas, definitions, concepts, examples, etc) in these notes is taken for teaching purposes, from several references:

- Numerical Analysis by R. L. Burden and J. D. Faires
- The Nature of Mathematical Modeling by N. Gershenfeld
- A First Course in Numerical methods by U. M. Ascher and C. Greif

These notes are only informally distributed and intended ONLY as study aid for the final exam of ETHZ students that were registered for the course Models, Algorithms and Data (MAD): Introduction to computing 2019. The notes have been checked, however they may still contain errors so use with care.

LECTURE 8 Numerical integration I: Rectangle, Trapezoidal and Simpson's Rule

We wish to evaluate a definite integral of the function $f(x)$ in the interval $x \in [a,b]$:



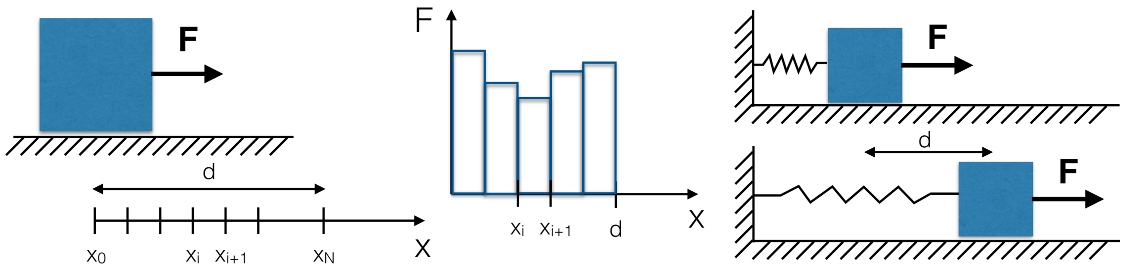
$$I = \int_a^b f(x)\,dx \tag{8.0.1}$$

Numerical integration is crucial in various situations:

- The integral cannot be solved analytically. Example: $I = \int_0^1 \sin(\cos(x))\,\mathrm{d}x$
- Function is only known for a set of *discrete points* as is the case for most experimental data. One can either find a functional form $f(x)$ (e.g. with the methods discussed earlier in this class) and integrate $f(x)$ analytically or one needs to use numerical integration.

Example 8.1: Work

We are pulling a brick on a surface with a force F.



We assume that there is no friction between the brick and the surface. If F is constant and we move the brick by a distance d, then the work we perform is equal to $A = Fd$. However, if F is not constant, i.e., $F = F(x)$ than we can make an approximation and assume that $F(x) = const.$ over a segment Δx (on interval $[x_i, x_{i+1}]$). The work we perform on this interval is $\Delta A_i = F_i \Delta x_i$

and the total work is

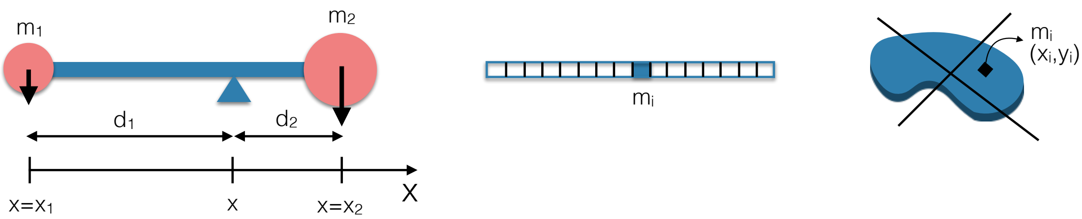
$$A = \sum_{i=0}^{N} \Delta A_i = \sum_{i=0}^{N} F(x_i)\Delta x_i,$$

where we have split the distance d into N intervals. If $N \to \infty$ and F is a continuous function we can write the total work as $A = \int_0^d F(x)\,dx$. If the brick is attached to the wall with a linear spring that has a spring constant k, we know that $F = kx$. Thus, the work we perform when we move the brick for a distance d is

$$A = \int_0^d kx\,dx = \frac{kd^2}{2}.$$

Example 8.2: Moments

Consider a beam that has at left and right ends attached a mass of mass m_1 and m_2, respectively. The question is how to find the position x of the support such that the beam is horizontal.



If we suppose that the beam has no weight, i.e., $m = 0$, then the balance of moments requires that $m_1(\bar{x} - x_1) = m_2(\bar{x} - x_2)$ which gives $\bar{x} = \frac{m_1 x_1 + m_2 x_2}{m_1 + m_2}$. If a beam is not massless, then we can split the beam into small segments and the center-of-mass position of the beam is given by

$$\bar{x} = \frac{\sum_{i=0}^{N} m_i x_i}{\sum_{i=0}^{N} m_i}.$$

Equivalently, we can write for the 2D plate $\bar{x}, \bar{y} = (\sum_{i=0}^{N} m_i(x_i, y_i))/(\sum_{i=0}^{N} m_i)$. Suppose that the density of the beam is not homogeneous, i.e., $\rho_i = m_i/\Delta x_i$ or $m_i = \rho_i \Delta x_i$. Then we can write the above equation as

$$\bar{x} = \frac{\sum_{i=0}^{N} \rho_i(x_i)\Delta x_i x_i}{\sum_{i=0}^{N} \rho_i(x_i)\Delta x_i} \xrightarrow{\text{N}\to\infty} \frac{\int \rho(x)x\,dx}{\int \rho(x)\,dx}.$$

8.1 Key idea

In numerical integration schemes, a common approach is to split the domain $[a, b]$ into N intervals $[x_i, x_{i+1}]$ of length $\Delta_i = x_{i+1} - x_i$. Here, we exploit the property of integrals

$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx, \tag{8.1.1}$$

i.e., an integral can then be evaluated as a sum of integrals over subdomains. Depending on the application, these points may be given a-priori or they are chosen for a given integration scheme. We may thus rewrite our integral as

$$I = \int_a^b f(x)\,dx = \sum_{i=0}^{N-1} \int_{x_i}^{x_{i+1}} f(x)\,dx \tag{8.1.2}$$

In a next step we approximate $f(x)$ within each interval with a simple function $p(x)$ that is easily integrable. The key idea of the numerical integration schemes is thus based on a "divide" and "conquer" approach. Where in step 1 we do a piecewise approximation of $f(x)$ and in step 2 we compute many such integrals exactly.

$$\boxed{I \approx \sum_{i=0}^{N-1} \int_{x_i}^{x_{i+1}} p_i(x)\,dx} \tag{8.1.3}$$

8.2 Numerical Quadrature

Different numerical integration schemes use different approximations to $f(x)$ in each interval $[x_i, x_{i+1}]$.

Rectangle Rule: We approximate $f(x)$ as constant using a single (left) point:

$$\boxed{I_{R_i} = f(x_i)\,\Delta_i} \tag{8.2.1}$$

Midpoint Rule: We approximate $f(x)$ as constant using the middle point:

$$\boxed{I_{M_i} = f(x_{i+1/2})\,\Delta_i = f\left(\frac{x_i + x_{i+1}}{2}\right)\Delta_i} \tag{8.2.2}$$

Trapezoidal Rule: We approximate $f(x)$ by a line (using 2 points):

$$\boxed{I_{T_i} = \frac{f(x_i) + f(x_{i+1})}{2}\Delta_i} \tag{8.2.3}$$

Simpson's Rule: We approximate $f(x)$ by a parabola (using 3 points):

$$I_{S_i} = \frac{f(x_i) + 4f((x_i + x_{i+1})/2) + f(x_{i+1})}{6}\Delta_i \tag{8.2.4}$$

Figure 8.1 shows an example of such approximations.

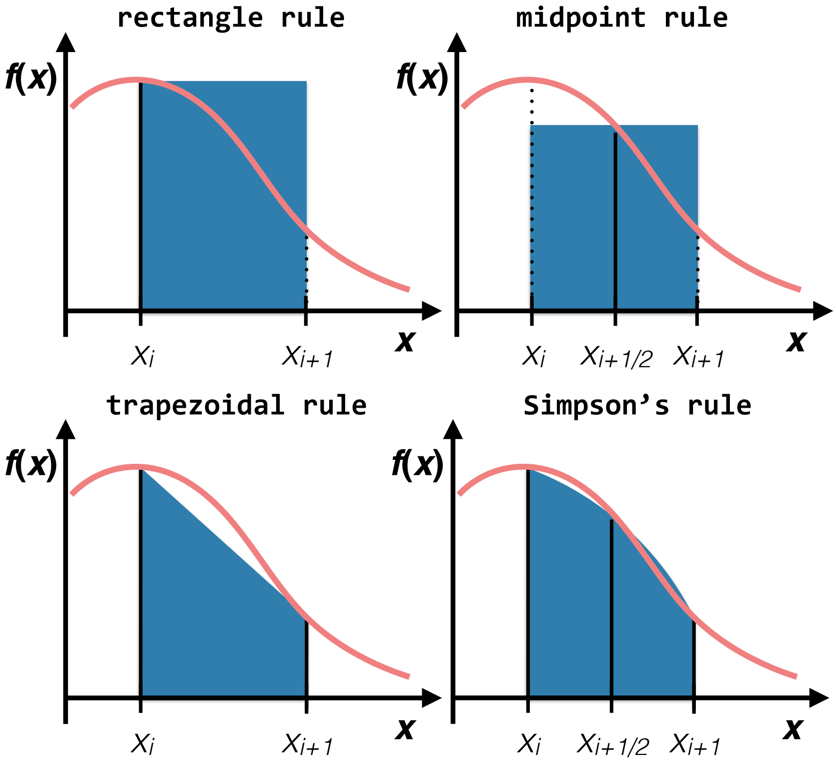


Figure 8.1: Example comparing the rectangle, midpoint, trapezoidal and the Simpson's rule for numerical integration.

Under the assumption of constant spacing $\Delta_i \equiv \Delta_x$ $(\forall i,\, i = 1,\ldots,N)$ we can now compute the total integral

$$\begin{aligned}
&\textbf{Rectangle Rule:} && I \approx \Delta_x \sum_{i=0}^{N-1} f(x_i),\\
&\textbf{Midpoint Rule:} && I \approx \Delta_x \sum_{i=0}^{N-1} f\left(\frac{x_i + x_{i+1}}{2}\right),\\
&\textbf{Trapezoidal Rule:} && I \approx \frac{\Delta_x}{2}\left(f(x_0) + 2\sum_{i=1}^{N-1} f(x_i) + f(x_N)\right),
\end{aligned} \tag{8.2.5}$$

For Simpson's approximation, we can rewrite Eq. (8.1.2) to sum over larger intervals $[x_i, x_{i+2}]$. In this way, we are evaluating $f(x)$ only at points x_i. If we assume that the total number of points $N+1$ is odd we obtain:

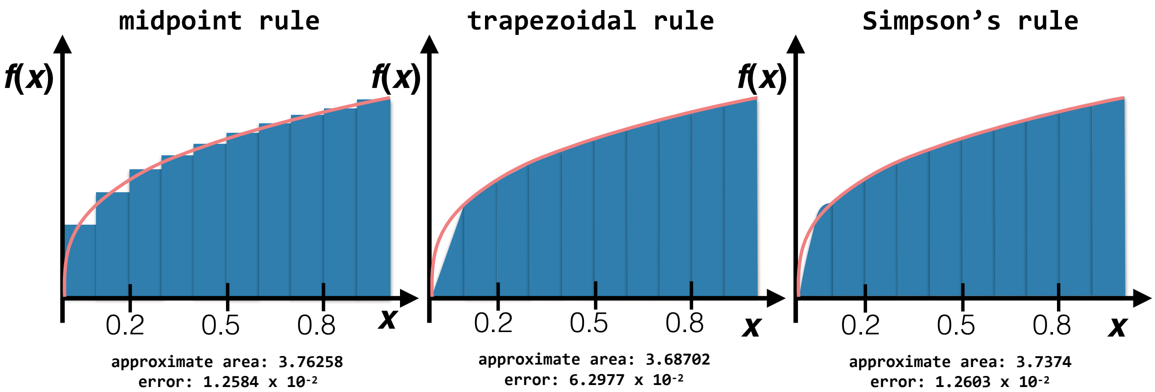
$$I = \int_a^b f(x)\,dx = \sum_{\substack{i=0 \\ i=\text{even}}}^{N-2} \int_{x_i}^{x_{i+2}} f(x)\,dx$$

$$\int_{x_i}^{x_{i+2}} f(x)\,dx \approx \frac{f(x_i) + 4f(x_{i+1}) + f(x_{i+2})}{3}\Delta_x$$

$$\textbf{Simpson's Rule:}\quad I \approx \frac{\Delta_x}{3}\left(f(x_0) + 4\sum_{\substack{i=1 \\ i=\text{odd}}}^{N-1} f(x_i) + 2\sum_{\substack{i=2 \\ i=\text{even}}}^{N-2} f(x_i) + f(x_N)\right). \tag{8.2.6}$$

Example 8.3: Composite Integration

We wish to evaluate $I = \int_0^1 5\sqrt[3]{x}\,dx$, i.e., the shaded area in Figure below.



To evaluate the integral we discretize the interval $[0,1]$ using $x_i = 0.1\,i$ $(i = 0,\ldots,10)$, i.e., we have $N = 10$ and $\Delta_x = 0.1$.

We thus see that all of the most fundamental methods write an integral as a weighted sum of f_i:

$$I = \int_a^b f(x)\,dx \approx \sum_{i=0}^{N} w_i f(x_i), \tag{8.2.7}$$

where w_i are the weights. For the trapezoidal rule for instance, we have $w_0 = w_N = \Delta_x/2$ and $w_i = \Delta_x$ for $i = 1,\ldots,N-1$.

8.2.1 Newton–Cotes formulas

A general way to derive quadrature rules is given by the Newton–Cotes formulas. To construct the approximate function $p_i(x)$ in Eq. (8.1.2) we use $M+1$ equidistant points in $[x_i, x_{i+1}]$ $(x_k = x_i + k\,h,$

$k = 0, \dots, M$) and Lagrange interpolation. The Lagrange interpolant through the points $(x_k, f(x_k))$ is given by

$$\begin{aligned} p_i(x) &= \sum_{k=0}^{M} f(x_k) l_k^M(x), \\ l_k^M(x) &= \frac{(x-x_0)(x-x_1)\dots(x-x_{k-1})(x-x_{k+1})\dots(x-x_M)}{(x_k-x_0)(x_k-x_1)\dots(x_k-x_{k-1})(x_k-x_{k+1})\dots(x_k-x_M)}. \end{aligned} \tag{8.2.8}$$

The integral I is then approximated as

$$I_i = \int_{x_i}^{x_{i+1}} f(x)\,dx \approx \int_{x_i}^{x_{i+1}} p_i(x)\,dx = \sum_{k=0}^{M} f(x_k) \int_{x_i}^{x_{i+1}} l_k^M(x)\,dx. \tag{8.2.9}$$

We rewrite this as

$$I_i \approx \Delta_i \sum_{k=0}^{M} C_k^M f(x_k), \text{ where } C_k^M = \frac{1}{\Delta_i} \int_{x_i}^{x_{i+1}} l_k^M(x)\,dx \tag{8.2.10}$$

Properties of C_k^M:

- Lagrange polynomials fit exactly a constant function: e.g. $f(x) = 1$ must be integrated exactly. In Eq. (8.2.10), we hence want to have

$$I_i = \int_{x_i}^{x_{i+1}} 1\,dx = (x_{i+1} - x_i) \sum_{k=0}^{M} C_k^M 1 \tag{8.2.11}$$

  and it follows that

$$\sum_{k=0}^{M} C_k^M = 1 \tag{8.2.12}$$

- For $x \in [x_i, x_{i+1}]$ we have that $l_{M-k}^M(x) = l_k^M(x_i + x_{i+1} - x)$. To see this, inserting $x_{M-j} = x_i + (M - j)h$ on the right hand side ($h = \frac{x_{i+1}-x_i}{M}$ and j some arbitrary number). After some algebra we find

$$l_k^M(x_i + x_{i+1} - x_{M-j}) = \cdots = l_k^M(x_j) = \delta_{jk} = l_{M-k}^M(x_{M-j})$$

  by definition of the Lagrange Polynomial. Since we did the calculation for an arbitrary j the above equality holds. If we now use this on our definition of the coefficient C_k^M and use a substitution of variables we find that

$$C_k^M = C_{M-k}^M \tag{8.2.13}$$

Example 8.4

Case: M=1

We choose $M = 1$ (i.e. 2 points $x_0 = x_i$, $x_1 = x_{i+1}$). The Lagrange polynomials are then given by

$$l_0^1(x) = \frac{x - x_1}{x_0 - x_1} = \frac{x_{i+1} - x}{x_{i+1} - x_i}, \quad l_1^1(x) = \frac{x - x_0}{x_1 - x_0} = \frac{x - x_i}{x_{i+1} - x_i}.$$

We integrate this and obtain:

$$C_0^1 = \frac{1}{\Delta_i}\int_{x_i}^{x_{i+1}} l_0^1(x)\,dx = \frac{1}{\Delta_i^2}\int_{x_i}^{x_{i+1}} (x_{i+1}-x)\,dx = \frac{1}{\Delta_i^2}\left(-\frac{1}{2}(x_{i+1}-x)^2\Big|_{x_i}^{x_{i+1}}\right) = \frac{1}{2},$$
$$C_1^1 = \frac{1}{\Delta_i}\int_{x_i}^{x_{i+1}} l_1^1(x)\,dx = \frac{1}{\Delta_i^2}\int_{x_i}^{x_{i+1}} (x-x_i)\,dx = \frac{1}{\Delta_i^2}\left(\frac{1}{2}(x-x_i)^2\Big|_{x_i}^{x_{i+1}}\right) = \frac{1}{2}.$$

If we insert $C_0^1 = C_1^1 = 1/2$ in Eq. (8.2.10), we see that we recover the trapezoidal rule:

$$I_i \approx \Delta_i \sum_{k=0}^{M} C_k^M f(x_k) = \Delta_i \frac{f(x_i)+f(x_{i+1})}{2}.$$

Case: M=2
For $M=2$ (i.e. 3 points $x_0 = x_i$, $x_1 = (x_i + x_{i+1})/2$, $x_2 = x_{i+1}$), we obtain Simpson's rule: $C_0^2 = 1/6$, $C_1^2 = 2/3$, $C_2^2 = 1/6$.

8.2.2 Error analysis

We would like to evaluate the error that with are making when we use different numerical integrations, that is we would like to find an upper bound or an approximation of

$$E_{\text{rule},i} = I_i - I_{\text{rule},i} \tag{8.2.14}$$

We will do so by using Taylor expansions to evaluate the error of the numerical integration schemes.

Midpoint rule: Consider Taylor's series around $x_{i+1/2} = (x_i + x_{i+1})/2$:

$$\begin{aligned} f(x) =& f(x_{i+1/2}) + \left(x - x_{i+1/2}\right) f'(x_{i+1/2}) + \frac{1}{2}\left(x - x_{i+1/2}\right)^2 f''(x_{i+1/2}) \\ &+ \frac{1}{6}\left(x - x_{i+1/2}\right)^3 f'''(x_{i+1/2}) + \cdots, \end{aligned} \tag{8.2.15}$$

We can use this expansion to evaluate I_i:

$$\begin{aligned} I_i &= \int_{x_i}^{x_{i+1}} f(x)\,dx \\ &= f(x_{i+1/2})\int_{x_i}^{x_{i+1}} dx + f'(x_{i+1/2})\int_{x_i}^{x_{i+1}} \left(x - x_{i+1/2}\right) dx \\ &\quad + \frac{1}{2} f''(x_{i+1/2})\int_{x_i}^{x_{i+1}} \left(x - x_{i+1/2}\right)^2 dx + \frac{1}{6} f'''(x_{i+1/2})\int_{x_i}^{x_{i+1}} \left(x - x_{i+1/2}\right)^3 dx + \cdots \\ &= \underbrace{f(x_{i+1/2})\,\Delta_i}_{I_{M_i}} + \underbrace{0 + \frac{1}{24} f''(x_{i+1/2})\Delta_i^3 + 0 + O(\Delta_i^5)}_{E_{M_i}} \end{aligned} \tag{8.2.16}$$

where $O(\Delta_i^5)$ includes all the higher order terms that we dropped from the Taylor series. So for the midpoint rule we have that:

$$\boxed{E_{M_i} = \frac{1}{24} f''(x_{i+1/2})\Delta_i^3 + O(\Delta_i^5) + \cdots} \tag{8.2.17}$$

For one interval the midpoint rule is third order accurate.

Trapezoidal rule: One can show that:

$$I_{T_i} = \frac{f(x_i) + f(x_{i+1})}{2}\Delta_i = \Delta_i \left(f(x_{i+1/2}) + \frac{1}{8} f''(x_{i+1/2})\Delta_i^2 + \cdots \right) = I_{M_i} + \frac{1}{8} f''(x_{i+1/2})\Delta_i^3 + \cdots$$

If we compare this with Eq. (8.2.17), we see that

$$\boxed{E_{T_i} = -\frac{1}{12} f''(x_{i+1/2})\Delta_i^3 + O(\Delta_i^5) + \cdots} \tag{8.2.18}$$

Simpson's rule: We can combine the midpoint and the trapezoidal rule. By comparing Eq. (8.2.4) with Eq. (8.2.2) and Eq. (8.2.3), we see that:

$$I_{S_i} = \frac{2}{3} I_{M_i} + \frac{1}{3} I_{T_i} \tag{8.2.19}$$

therefore

$$\boxed{E_{S_i} = O(\Delta_i^5) + \cdots}\,. \tag{8.2.20}$$

We remark that the above error estimates are only valid for the local approximation. When considering the whole domain (with a constant spacing $\Delta_x = \Delta_i$ $(\forall i)$) the error is reduced to second order for the midpoint and trapezoidal rule and fourth order for Simpson's rule. For the midpoint rule for instance we get (using Eq. (8.1.2)):

$$\begin{aligned} \sum_{i=0}^{N-1} I_{M_i} &= \sum_{i=0}^{N-1} \left(I_i - \frac{1}{24} f''(x_{i+1/2})\Delta_x^3 + O(\Delta_x^5) + \cdots \right), \\ \left| \sum_{i=0}^{N-1} I_{M_i} - I \right| &< N \frac{1}{24} \max_i (|f''(x_{i+1/2})|)\Delta_x^3 + N O(\Delta_x^5) + \cdots, \\ &= \frac{b-a}{24} \max_i (|f''(x_{i+1/2})|)\Delta_x^2 + O(\Delta_x^4) + \cdots, \end{aligned} \tag{8.2.21}$$

where we used that $N = (b-a)/\Delta_x$.

Exam checklist

After this class, you should understand the following points regarding numerical integration:

- When to use numerical integration.
- How to do numerical integration with the rectangle, midpoint, trapezoidal and Simpson's rule.

- How to estimate errors of a numerical integration scheme.