Chapter 6
Numerical integration

Let us consider numerical integration such as

$$\int_0^1 f(x)\mathrm{d}x \approx \sum_{i=0}^{n} w_i f_i,$$

where w_i are coefficients and $f_i = f(x_i)$ for x_i $(i = 0, 1, \ldots, n)$.

Trapezoid rule

Let us consider

$$\int_0^1 \mathrm{e}^{-x^2}\mathrm{d}x = 0.7468\ldots,$$

with uniform points:

$$x_i = ih, \quad h = \frac{1}{n}, \quad i = 0, 1, \ldots, n.$$

Probably the most naive implementation is the (right-hand) Riemann sum:

$$\int f(x)\mathrm{d}x \approx R(h) = f_1 h + f_2 h + \cdots + f_n h = (f_1 + f_2 + \cdots + f_n)\,h.$$

In the present example, we have

h	$R(h)$	error	error$/h$	error$/h^2$
1	0.3679	0.3789	0.3789	0.3789
0.5	0.5733	0.1735	0.3470	0.6939
0.25	0.6640	0.0829	0.3314	1.3257
0.125	0.7064	0.0405	0.3237	2.5898

We see that the error is order h and the method is first-order accurate.

The trapezoid rule is given by

Fall 2013 Math 471 Sec 2
Introduction to Numerical Methods

Manabu Machida (University of Michigan)

$$\int f(x)\mathrm{d}x \approx T(h) = \frac{1}{2}(f_0+f_1)h + \frac{1}{2}(f_1+f_2)h + \cdots + \frac{1}{2}(f_{n-1}+f_n)h$$
$$= \left(\frac{1}{2}f_0 + f_1 + f_2 + \cdots + f_{n-1} + \frac{1}{2}f_n\right)h.$$

We obtain

h	$T(h)$	error	error$/h$	error$/h^2$
1	0.6839	0.0629	0.0629	0.0629
0.5	0.7314	0.0155	0.0309	0.0618
0.25	0.7430	0.0038	0.0154	0.0614
0.125	0.7459	0.0010	0.0077	0.0613

We see that the error is $O(h^2)$ and it is second-order accurate.

It is a natural question how we can obtain more accurate integration formulae.

1. piecewise quadratic interpolant (Simpson's rule)[1]
2. cubic spline interpolant
3. non-uniform points (e.g., Chebyshev)
4. extrapolation

We will explore the last option in the next section.

Richardson extrapolation (Romberg's method)

We first define $R_0(h)$ as[2]

---

[1] For $n = 2m$ uniform points $x_0, x_1, \ldots, x_{2m}$, Simpson's rule is given as follows.

$$\int f(x)\mathrm{d}x \approx \frac{2h}{6}(f_0+4f_1+f_2) + \frac{2h}{6}(f_2+4f_3+f_4) + \frac{2h}{6}(f_4+4f_5+f_6) + \cdots + \frac{2h}{6}(f_{2m-2}+4f_{2m-1}+f_{2m})$$
$$= \frac{h}{3}\left(f_0 + 4\sum_{j=1}^{m} f_{2j-1} + 2\sum_{j=1}^{m-1} f_{2j} + f_{2m}\right).$$

We note that using the Lagrange form

$$\int_{x_0}^{x_2} p_2(x)\mathrm{d}x = \int_{x_0}^{x_2}\left(f_0\frac{x-x_1}{x_0-x_1}\frac{x-x_2}{x_0-x_2} + f_1\frac{x-x_0}{x_1-x_0}\frac{x-x_2}{x_1-x_2} + f_2\frac{x-x_0}{x_2-x_0}\frac{x-x_1}{x_2-x_1}\right)\mathrm{d}x = \cdots = \frac{h}{3}(f_0+4f_1+f_2).$$

[2] We focus on the interval $[x_{i-1}, x_i]$ and define $m_i = (x_{i-1}+x_i)/2$. We note that the Taylor series for f around the midpoint m_i is given by $f(x) = f(m_i) + (x-m_i)f'(m_i) + \frac{1}{2}(x-m_i)^2 f''(m_i) + \frac{1}{3!}(x-m_i)^3 f'''(m_i) + \frac{1}{4!}(x-m_i)^4 f^{(4)}(m_i) + \cdots$. Hence,

$$R_0(h) = T(h) = \int_0^1 f(x)\mathrm{d}x + c_2h^2 + c_4h^4 + c_6h^6 + \cdots.$$

Note that $T(h)$ is second-order accurate. We consider

$$\begin{aligned} R_0(2h) = T(2h) &= \int_0^1 f(x)\mathrm{d}x + c_2(2h)^2 + c_4(2h)^4 + c_6(2h)^6 + \cdots \\ &= \int_0^1 f(x)\mathrm{d}x + 4c_2h^2 + 16c_4h^4 + 64c_6h^6 + \cdots. \end{aligned}$$

By subtraction we obtain

$$4R_0(h) - R_0(2h) = 3\int_0^1 f(x)\mathrm{d}x - 3\cdot 4c_4h^4 - 3\cdot 20c_6h^6 + \cdots.$$

We define

$$R_1(h) = \frac{1}{3}\left(4R_0(h) - R_0(2h)\right) = R_0(h) + \frac{1}{3}\left(R_0(h) - R_0(2h)\right).$$

We see that $R_1(h)$ is 4th order accurate. That is,

$$R_1(h) = \int_0^1 f(x)\mathrm{d}x + \tilde{c}_4h^4 + \tilde{c}_6h^6 + \cdots.$$

We further consider

$$\begin{aligned} R_1(2h) &= \int_0^1 f(x)\mathrm{d}x + \tilde{c}_4(2h)^4 + \tilde{c}_6(2h)^6 + \cdots \\ &= \int_0^1 f(x)\mathrm{d}x + 16\tilde{c}_4h^4 + 64\tilde{c}_6h^6 + \cdots. \end{aligned}$$

$$\begin{aligned} \frac{h}{2}\left(f(x_{i-1}) + f(x_i)\right) - \int_{x_{i-1}}^{x_i} f(x)\mathrm{d}x &= \frac{h}{2}\left(2f(m_i) + 0 + \frac{f''(m_i)}{4} + 0 + \frac{f^{(4)}(m_i)}{192}h^4 + \cdots\right) \\ &\quad - \left(hf(m_i) + 0 + \frac{f''(m_i)}{24}h^3 + 0 + \frac{f^{(4)}(m_i)}{1920}h^5 + \cdots\right) \\ &= \frac{f''(m_i)}{12}h^3 + \frac{f^{(4)}(m_i)}{480}h^5 + \cdots. \end{aligned}$$

Thus,

$$\begin{aligned} T(h) - \int_0^1 f(x)\mathrm{d}x &= \sum_{i=1}^{n}\left(\frac{f''(m_i)}{12}h^3 + \frac{f^{(4)}(m_i)}{480}h^5 + \cdots\right) \\ &= \frac{1}{12}\frac{\sum_{i=1}^n f''(m_i)}{n}nh^3 + \frac{1}{480}\frac{\sum_{i=1}^n f^{(4)}(m_i)}{n}nh^5 + \cdots \\ &= \frac{1}{12}\left[f''(m_i)\right]_{\mathrm{ave}}h^2 + \frac{1}{480}\left[f^{(4)}(m_i)\right]_{\mathrm{ave}}h^4 + \cdots, \end{aligned}$$

where we used $n = 1/h$. Note that only even powers appear in the error.

By subtraction we have

$$16R_1(h) - R_1(2h) = 15\int_0^1 f(x)\mathrm{d}x - 15\cdot\frac{16}{5}\tilde{c}_6 h^6 + \cdots.$$

We introduce

$$R_2(h) = \frac{1}{15}\left(16R_1(h) - R_1(2h)\right) = R_1(h) + \frac{1}{15}\left(R_1(h) - R_1(2h)\right).$$

Thus $R_2(h)$ is 6th order accurate. By repeating this procedure we can make better formulae. We obtain

$$R_k(h) = \frac{1}{4^k - 1}\left(4^k R_{k-1}(h) - R_{k-1}(2h)\right), \quad k = 1, 2, \ldots.$$

For example,

$$R_3(h) = \frac{1}{63}\left(64R_2(h) - R_2(2h)\right).$$

Example 1. Let us consider

$$\int_0^1 \mathrm{e}^{-x^2}\mathrm{d}x = 0.7468\ldots,$$

with uniform points:

$$x_i = ih, \quad h = \frac{1}{n}, \quad i = 0, 1, \ldots, n.$$

By Richardson extrapolation we obtain the following numerical results.

h	$R_0(h)$	$R_1(h)$	$R_2(h)$	$R_3(h)$
1.0	0.683940	–	–	–
0.5	0.731370	0.7471800	–	–
0.25	0.742984	0.7468553	0.7468336	–
0.125	0.745866	0.7468266	0.7468246	0.7468244

If we go down a column decreasing h, then we have a fixed order of accuracy. If we go across a row with a fixed h, then the order of accuracy increases.

Extrapolation can be applied to any numerical approximation if the error has an expansion in powers of h. Here we used the Richardson extrapolation for numerical integration (Romberg's method).

Orthogonal polynomials

Let us begin by recalling

$$\mathbf{x} \cdot \mathbf{y} = \mathbf{x}^T \mathbf{y} = \sum_{i=1}^{n} x_i y_i.$$

Definition 1. The inner product of two functions f, g on the interval $[-1, 1]$ is defined as

$$\langle f, g \rangle = \int_{-1}^{1} f(x) g(x) \mathrm{d}x.$$

Two functions f, g are said to be orthogonal if $\langle f, g \rangle = 0$. Also, $\|f\| = \sqrt{\langle f, f \rangle}$ is called a norm of f.

We have the following properties.

1. $\langle f, f \rangle \geq 0, \quad \|f\| = 0 \Leftrightarrow f = 0$
2. $\langle f, \alpha g + h \rangle = \alpha \langle f, g \rangle + \langle f, h \rangle$

Example 2. The functions $\sin(\pi x)$ and $\cos(\pi x)$ are orthogonal because

$$\langle \sin(\pi x), \cos(\pi x) \rangle = \int_{-1}^{1} \sin(\pi x) \cos(\pi x) \mathrm{d}x = \frac{1}{2} \int_{-1}^{1} \sin(2\pi x) \mathrm{d}x = 0.$$

The functions 1 and x are orthogonal because

$$\langle 1, x \rangle = \int_{-1}^{1} 1 \cdot x \mathrm{d}x = 0.$$

The functions 1 and x^2 are not orthogonal because

$$\langle 1, x^2 \rangle = \int_{-1}^{1} 1 \cdot x^2 \mathrm{d}x = \frac{2}{3} \neq 0.$$

From the above examples, we see that if $f(x)$ is even and $g(x)$ is odd (or vice versa), then f and g are orthogonal (on $[-1, 1]$).

Starting from $\{1, x, x^2, \ldots\}$, the Gram-Schmidt process yields a set of orthogonal polynomials $\{P_0(x), P_1(x), P_2(x), \ldots\}$ which are called the Legendre polynomials[3].

---

[3] A polynomial is said to be monic if its leading coefficient is $+1$. Legendre polynomials in this section are monic. However, often $P_n(1) = 1$ is imposed and we get

$$P_0(x) = 1, \quad P_1(x) = x, \quad P_2(x) = \frac{1}{2}(3x^2 - 1), \quad P_3(x) = \frac{1}{2}(5x^3 - 3x), \ldots.$$

By the way, they satisfy the three-term recurrence relation $(n+1)P_{n+1}(x) = (2n+1)xP_n(x) - nP_{n-1}(x)$.

$$\begin{aligned}
P_0(x) &= 1,\\
P_1(x) &= x - \frac{\langle x, P_0\rangle}{\|P_0\|^2}P_0 = x,\\
P_2(x) &= x^2 - \frac{\langle x^2, P_0\rangle}{\|P_0\|^2}P_0 - \frac{\langle x^2, P_1\rangle}{\|P_1\|^2}P_1 = x^2 - \frac{1}{3},\\
P_3(x) &= x^3 - \frac{\langle x^3, P_0\rangle}{\|P_0\|^2}P_0 - \frac{\langle x^3, P_1\rangle}{\|P_1\|^2}P_1 - \frac{\langle x^3, P_2\rangle}{\|P_2\|^2}P_2 = x^3 - \frac{3}{5}x,
\end{aligned}$$

and so on. Note that they are orthogonal:

$$\langle P_i, P_j\rangle = 0, \quad i \neq j.$$

Gaussian quadrature

Here we consider a numerical integral

$$\int_{-1}^{1} f(x)\mathrm{d}x \approx \sum_{i=1}^{n} w_i f(x_i),$$

where w_i, x_i are chosen such that

$$\int_{-1}^{1} p(x)\mathrm{d}x = \sum_{i=1}^{n} w_i p(x_i) \qquad \text{(exact!)} \tag{6.1}$$

for any polynomial $p(x)$ of degree $\leq 2n-1$.

Let us first try a brute force method.

Example 3. When $n = 1$, we can write $p(x)$ as

$$p(x) = a_0 + a_1 x,$$

where a_0, a_1 are constants. Then (6.1) is written as

$$\int_{-1}^{1} (a_0 + a_1 x)\mathrm{d}x = w_1 f(x_1).$$

Here,

$$\text{LHS} = 2a_0 + 0, \qquad \text{RHS} = w_1(a_0 + a_1 x_1).$$

Therefore, we obtain

$$w_1 = 2, \qquad x_1 = 0.$$

The one-point Gaussian quadrature is obtained as

$$\int_{-1}^{1} f(x)\mathrm{d}x \approx 2f(0).$$

Note that 0 is the midpoint between -1 and 1.

Example 4. For $n=2$, we can write

$$p(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3.$$

We have

$$\begin{aligned}
\text{LHS of (6.1)} &= \int_{-1}^{1} (a_0 + a_1 x + a_2 x^2 + a_3 x^3) \mathrm{d}x \\
&= 2a_0 + \frac{2}{3} a_2, \\
\text{RHS of (6.1)} &= w_1 (a_0 + a_1 x_1 + a_2 x_1^2 + a_3 x_1^3) + w_2 (a_0 + a_1 x_2 + a_2 x_2^2 + a_3 x_2^3) \\
&= (w_1 + w_2) a_0 + (w_1 x_1 + w_2 x_2) a_1 + (w_1 x_1^2 + w_2 x_2^2) a_2 + (w_1 x_1^3 + w_2 x_2^3) a_3.
\end{aligned}$$

Note that we have four unknowns w_i, x_i and four equations.

$$w_1 + w_2 = 2 \quad \Rightarrow \quad w_2 = 2 - w_1.$$

$$w_1 x_1 + w_2 x_2 = 0 \quad \Rightarrow \quad x_2 = \frac{w_1}{w_1 - 2} x_1.$$

$$w_1 x_1^3 + w_2 x_2^3 = 0 \quad \Rightarrow \quad w_1 x_1^3 + (2 - w_1) \frac{w_1^3}{(w_1 - 2)^3} x_1^3 = 0 \quad \Rightarrow \quad w_1 = 1.$$

$$w_1 = 1 \quad \Rightarrow \quad w_2 = 1, \quad x_2 = -x_1.$$

$$w_1 x_1^2 + w_2 x_2^2 = \frac{2}{3} \quad \Rightarrow \quad 2x_1^2 = \frac{2}{3} \quad \Rightarrow \quad x_1 = \pm \frac{1}{\sqrt{3}}.$$

Thus the two-point Gaussian quadrature is obtained as

$$\int_{-1}^{1} f(x) \mathrm{d}x \approx f(-\frac{1}{\sqrt{3}}) + f(\frac{1}{\sqrt{3}}).$$

How can we determine w_i, x_i in general? We begin by the following theorem.

Theorem 1. *$P_n(x)$ ($n \geq 1$) has n distinct roots $x_1, \ldots, x_n$ on $(-1,1)$.*

Proof. Using the orthogonality relation for Legendre polynomials, we have $\int_{-1}^{1} P_n(x) \mathrm{d}x = \langle P_n, P_0 \rangle = 0$. Hence $P_n(x)$ changes sign at least once on $(-1,1)$. We assume that $P_n(x)$ changes sign j ($1 \leq j \leq n$) times at $x_1, \ldots, x_j$ on $(-1,1)$. The polynomial $q(x) = (x - x_1)(x - x_2) \cdots (x - x_j)$ of degree j changes sign at $x_1, \ldots, x_n$. This implies that $P_n(x)$ and $q(x)$ have the same signs for all $x \in (-1,1)$ or have the opposite signs for all x. In either case, $\langle P_n, q \rangle = \int_{-1}^{1} P_n(x) q(x) \mathrm{d}x \neq 0$. Thus the degree of $q(x)$ is $\geq n$ because we can write $q(x) = \sum_{i=0}^{j} c_i P_i(x)$ with some c_i and $\langle P_n, q \rangle = \sum_{i=0}^{j} c_i \langle P_n, P_i \rangle = 0$ if $j < n$. However, $j \leq n$. Therefore we conclude $j = n$. That is, $P_n(x)$ has n distinct roots.

Let us consider how we can determine w_i, x_i ($1 \leq i \leq n$). First we choose $x_1, \ldots, x_n$ from the roots of $P_n(x)$. To determine w_i we consider the following two cases.

Case 1:

Let $p(x)$ be a polynomial of degree $\leq n-1$. Using the Lagrange form, we can write

$$p(x) = \sum_{i=1}^{n} p(x_i) L_i(x),$$

where $L_i(x)$ are Lagrange interpolating polynomials:

$$L_i(x) = \prod_{\substack{j=1 \\ j \neq i}}^{n} \frac{x - x_j}{x_i - x_j}.$$

We then obtain

$$\int_{-1}^{1} p(x) \mathrm{d}x = \sum_{i=1}^{n} p(x_i) \int_{-1}^{1} L_i(x) \mathrm{d}x$$

By setting

$$w_i = \int_{-1}^{1} L_i(x) \mathrm{d}x, \tag{6.2}$$

we obtain (6.1).

Case 2:

Let $p(x)$ be a polynomial of degree $\leq 2n-1$. We can express p as

$$p(x) = q(x) P_n(x) + r(x),$$

where the quotient $q(x)$ and the remainder $r(x)$ are polynomials of degree $\leq n-1$. We have

$$\int_{-1}^{1} p(x) \mathrm{d}x = \int_{-1}^{1} q(x) P_n(x) \mathrm{d}x + \int_{-1}^{1} r(x) \mathrm{d}x = \langle q, P_n \rangle + \int_{-1}^{1} r(x) \mathrm{d}x = \int_{-1}^{1} r(x) \mathrm{d}x,$$

where we used the fact that by writing $q(x) = \sum_{i=0}^{n-1} c_i P_i(x)$ with some c_i, we get

$$\langle q, P_n \rangle = \sum_{i=0}^{n-1} c_i \langle P_i, P_n \rangle = 0.$$

Since $r(x)$ is a polynomial of degree $\leq n-1$, by Case 1, we have

$$\int_{-1}^{1} r(x) \mathrm{d}x = \sum_{i=1}^{n} w_i r(x_i).$$

We note that since $P_n(x_i) = 0$ $(1 \leq i \leq n)$,

$$p(x_i) = q(x_i)P_n(x_i) + r(x_i) = r(x_i).$$

Therefore,

$$\int_{-1}^{1} p(x)\mathrm{d}x = \sum_{i=1}^{n} w_i p(x_i),$$

where w_i are given in (6.2).

By Gaussian quadrature, we have

$$\int_{-1}^{1} f(x)\mathrm{d}x \approx \sum_{i=1}^{n} w_i f(x_i), \quad w_i = \int_{-1}^{1} L_i(x)\mathrm{d}x, \quad P_n(x_i) = 0, \quad i = 1, 2, \ldots, n,$$

where $L_i(x)$ are Lagrange interpolating polynomials.

We don't have to calculate w_i, x_i for every $f(x)$. We can prepare the following table.

n	x_i	w_i
1	0	2
2	$-\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}$	1
3	$-\sqrt{\frac{3}{5}}, 0, \sqrt{\frac{3}{5}}$	$\frac{5}{9}, \frac{8}{9}, \frac{5}{9}$
4	$-\sqrt{\frac{3+2\sqrt{6/5}}{7}}, -\sqrt{\frac{3-2\sqrt{6/5}}{7}}, \sqrt{\frac{3-2\sqrt{6/5}}{7}}, \sqrt{\frac{3+2\sqrt{6/5}}{7}}$	$\frac{18-\sqrt{30}}{36}, \frac{18+\sqrt{30}}{36}, \frac{18+\sqrt{30}}{36}, \frac{18-\sqrt{30}}{36}$
$\vdots$	$\vdots$	$\vdots$

Example 5. Let us compute $\int_0^1 \mathrm{e}^{-x^2}\mathrm{d}x$ by Gaussian quadrature. By changing the variable as $t = 2x - 1$, we have

$$\int_0^1 \mathrm{e}^{-x^2}\mathrm{d}x = \int_{-1}^{1} \exp\left[-\left(\frac{t+1}{2}\right)^2\right]\frac{\mathrm{d}t}{2} = 0.746824\ldots.$$

We compute $G_n = \sum_{i=1}^{n} w_i f(x_i)$ for different n. Recall $T(0.125) = 0.745866$ ($n = 8$) with the trapezoid rule. Gaussian quadrature is more accurate than the trapezoid rule.

n	G_n
1	0.778801
2	0.746595
3	0.746815
4	0.746824

Theorem 2. *The error for Gaussian quadrature is given as follows.*

$$\int_{-1}^{1} f(x)\mathrm{d}x - \sum_{i=1}^{n} w_i f(x_i) = \frac{\alpha_n}{a_n^2 (2n)!} f^{(2n)}(\xi),$$

where $\alpha_n = \int_{-1}^{1} P_n^2(x)\mathrm{d}x$*,* a_n *is the leading coefficient of* $P_n(x)$*, and* $\xi \in [-1,1]$.

Proof. See Exercises 31 and 32 in Section 6.6 of the textbook.

As the final comment, we note that Gaussian quadrature can be extended to other orthogonal polynomials such as Laguerre polynomials, Hermite polynomials, and Chebyshev polynomials.