Numerical analysis: numerical integration

Dr. Dmitry Batenkov

April 18, 2021

1 Approximating the integral by interpolating polynomials

We would like to compute approximation to $I(f)=\int_a^b f\left(x\right)dx$,by using values of $f\left(x\right)$ at a finite number of points $\{x_0,\ldots,x_n\}$. This need arises in the following scenarios:

- It is impossible to compute the indefinite integral
- $f\left(x\right)$ is given by its samples and not an explicit formula.
- Other "tricks" don't work

One common approach is to approximate f by its interpolation polynomial, and then compute the integral of this polynomial:

$$f\left(x\right)=P_n\left(x\right)+e\left(x\right)=P_n\left(x\right)+f\left[x_0,\ldots,x_n,x\right]\underbrace{q_{n+1}\left(x\right)}_{\prod_{k=0}^n(x-x_k)}$$

$$\int_a^b f(x)dx=\int_a^b P_n(x)dx+\int_a^b e(x)dx=\underbrace{\int_a^b P_n(x)dx}_{\text{the approximation}}+\underbrace{\int_a^b f\left[x_0,x_1,\ldots,x_n,x\right]q_{n+1}\left(x\right)dx}_{E(I(f))} \tag{1}$$

Let us now consider different important special cases where the error formula can be simplified.

<u>Case A:</u> $q_{n+1}\left(x\right)$ has constant sign in $[a,b]$. We can use the following result (without proof):

Claim 1. Let f,g be continuous in $[a,b]$. Without loss of generality, $h\left(x\right)\geq 0$ for $x\in[a,b]$. Then $\exists c\in[a,b]$ s.t.

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

Corollary 1. *If q_{n+1} has constant sign in $[a,b]$ then $\exists c,\xi\in[a,b]$ s.t. (recall Claim 5 from Lectures 1-3)*

$$E\left(I\left(f\right)\right)=f\left[x_0,x_1,\ldots,x_n,c\right]\int_a^b q_{n+1}(x)dx=\frac{f^{(n+1)}(\xi)}{(n+1)!}\int_a^b q_{n+1}(x)dx. \tag{2}$$

Putting $h=b-a$ we get $\left|q_{n+1}\left(x\right)\right|\leq Ch^{n+1}$ and therefore $\left|E\left(I\left(f\right)\right)\right|\leq C_1h^{n+2}$.

Example 1 (Rectangle rule)**.** Consider $P_0(x) = f(a)$, then

$$\int_a^b f(x)dx = \int_a^b P_0(x)dx + \int_a^b f[a,x]\psi(x)dx = \underbrace{(b-a)f(a)}_{\text{Rectangle Integration Rule}} + \underbrace{\int_a^b f[a,x](x-a)dx}_{\text{error}}.$$

Since $q_1(x) = (x-a) \geq 0$ we are in <u>**Case A**</u> and so (as before, $h = b-a$)

$$E(I(f)) = f[a,c]\int_a^b (x-a)dx = f'(\xi)\left[\frac{(x-a)^2}{2}\right]_a^b = \frac{(b-a)^2}{2}f'(\xi) = \frac{h^2}{2}f'(\xi).$$

So we got a scheme of <u>approximation order=2.</u>

<u>**Case B:**</u> $\int_a^b q_{n+1}(x)dx = 0$. In this case for every x_{n+1}

$$\begin{aligned}
f[x_0,x_1,\ldots,x_n,x_{n+1},x] &= \frac{f[x_0,x_1,\ldots,x_n,x] - f[x_0,x_1,\ldots,x_n,x_{n+1}]}{x-x_{n+1}} \\
f[x_0,x_1,\ldots,x_n,x] &= f[x_0,x_1,\ldots,x_n,x_{n+1}] + f[x_0,x_1,\ldots,x_n,x_{n+1},x](x-x_{n+1}) \\
E(I(f)) &= \int_a^b f[x_0,x_1,\ldots,x_n,x]\,q_{n+1}(x)\,dx \\
&= \int_a^b \{f[x_0,x_1,\ldots,x_n,x_{n+1}] + f[x_0,x_1,\ldots,x_n,x_{n+1},x](x-x_{n+1})\}\,q_{n+1}(x)\,dx \\
&= \int_a^b f[x_0,x_1,\ldots,x_n,x_{n+1},x](x-x_{n+1})\,q_{n+1}(x)\,dx.
\end{aligned}$$

Now if we can choose x_{n+1} s.t. $(x-x_{n+1})\,q_{n+1}(x)$ has constant sign in $[a,b]$, then as in Corollary 1 above we get further simplification:

$$E(I(f)) = f[x_0,x_1,\ldots,x_n,x_{n+1},c]\int_a^b (x-x_{n+1})\,q_{n+1}(x)dx = \frac{f^{(n+2)}(\xi)}{(n+2)!}\int_a^b (x-x_{n+1})\,q_{n+1}(x)dx. \tag{3}$$

In a similar fashion, for $h = b-a$ we get $|(x-x_{n+1})\,q_{n+1}(x)| \leq Ch^{n+2}$ and $|E(I(f))| \leq C_1 h^{n+3}$. So the scheme has approximation order $n+3$, higher than in <u>**Case A.**</u>

Example 2. Still order $n = 0$ but computing $P_0(x)$ using the point $x_0 = \frac{a+b}{2}$, i.e. $P_0(x) = f\left(\frac{a+b}{2}\right)$. We get

$$\begin{aligned}
\int_a^b f(x)dx &= \int_a^b P_0(x)dx + \int_a^b f\left[\frac{a+b}{2},x\right]q_1(x)dx. \\
&= \underbrace{(b-a)f\left(\frac{a+b}{2}\right)}_{\text{Midpoint Integration Rule}} + \underbrace{\int_a^b f\left[\frac{a+b}{2},x\right]\left(x-\frac{a+b}{2}\right)dx}_{\text{error}}.
\end{aligned}$$

Since $\int_a^b q_1(x)dx = \int_a^b \left(x - \frac{a+b}{2}\right) dx = 0$ we are in **Case B**. Choose $x_1 = x_0$, then $(x - x_1)\, q_1(x) = (x - x_0)^2$ has constant sign, therefore

$$\begin{aligned} E(I(f)) &= \int_a^b f\left[\frac{a+b}{2}, \frac{a+b}{2}, x\right] \left(x - \frac{a+b}{2}\right)^2 dx \\ &= f\left[\frac{a+b}{2}, \frac{a+b}{2}, c\right] \int_a^b \left(x - \frac{a+b}{2}\right)^2 dx \\ &= \frac{f^{(2)}(\xi)}{2!} \left[\frac{\left(x - \frac{a+b}{2}\right)^3}{3}\right]_a^b \\ &= \frac{f^{(2)}(\xi)(b-a)^3}{24}. \end{aligned}$$

With $h = b - a$ we get $E(I(f)) = \frac{h^3}{24} f''(\xi)$, i.e. a scheme of <u>approximation order=3,</u> **<u>by using the same number of points as in the Rectangle rule!</u>**

Example 3 (Trapezoidal Rule). Now we use $n = 1$ with endpoints $x_0 = a, x_1 = b$.

$$\begin{aligned} P_1(x) &= f(a) + f[a,b](x-a) = f(a) + \frac{f(b) - f(a)}{b-a}(x-a) \\ \int_a^b P_1(x)\, dx &= f(a)(b-a) + \frac{f(b) - f(a)}{2}(b-a) = \frac{f(a) + f(b)}{2}(b-a) \qquad \Leftarrow \text{Area of a trapezoid!} \\ \int_a^b f(x)\, dx &= \int_a^b P_1(x)dx + \int_a^b f[a,b,x] q_2(x) dx \\ &= \underbrace{\frac{b-a}{2}(f(a) + f(b))}_{\text{Trapezoidal Integration Rule}} + \underbrace{\int_a^b f[a,b,x](x-a)(x-b)dx}_{\text{error}}. \end{aligned}$$

Error computation: $q_2(x) = (x-a)(x-b) \leq 0$ so we are in **Case A**: $h = b - a$

$$E(I(f)) = \frac{f''(\xi)}{2} \int_a^b q_2(x)dx = -\frac{f''(\xi)}{12} h^3.$$

Thus <u>approximation order=3.</u>

2 Method of undetermined coefficients

Recall the definition of algebraic degree of exactness from Lecture 4. An alternative derivation seeks a scheme of the form

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

which should have algebraic degree of exactness at least n. Thus it should be exact for $f = 1, x, \ldots, x^n$. This implies that $\{A_0, \ldots, A_n\}$ should satisfy the linear system of equations

$$\begin{array}{l} I(1)=\int_a^b 1dx=b-a=A_0\cdot 1+A_1\cdot 1+\ldots+A_n\cdot 1 \\ I(x)=\int_a^b xdx=\frac{b^2}{2}-\frac{a^2}{2}=A_0x_0+A_1x_1+\ldots+A_nx_n \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots.. \\ I\left(x^n\right)=\int_a^b x^n dx=\frac{b^{n+1}-a^{n+1}}{n+1}=A_0x_0^n+A_1x_1^n+\ldots+A_nx_n^n \end{array}.$$

Indeed, if $I(f) = \tilde{I}(f)$ for $f = 1, x, \ldots, x^n$ then $I(f) = \tilde{I}(f)$ for all $f \in \Pi_n$ by linearity (both $I, \tilde{I}$ are linear operators). Alternatively, we can construct the system as above for any linearly independent set $\{p_1, p_2, \ldots, p_n\}$ which spans Π_n.

Claim 2. The schemes thus obtained are equivalent to the ones derived via explicit form of P_n, and so the error formula (1) holds.

Proof. From the form of the system of equations above, there is always a unique solution $\{A_0, \ldots, A_n\}$ for every choice of pairwise distinct $\{x_0, \ldots, x_n\}$ (why?) On the other hand, integrating the Lagrange form $P_n(x) = \sum_{i=0}^n f(x_i)\,\ell_i(x)$ we get

$$\hat{I}(f)=\int_a^b P_n(x)\,dx=\sum_{i=0}^n f(x_i)\underbrace{\int_a^b \ell_i(x)\,dx}_{:=B_i},$$

and this equation should also be exact for $\{1, x, \ldots, x^n\}$. So $\{A_i\}_{i=0}^n$ and $\{B_i\}_{i=0}^n$ satisfy the same system of equations, thus $A_i = B_i$. □

Example 4 (Simpson's Rule)**.** $n = 2$, looking for coefficients A, B, C such that $\tilde{I}(f) = Af(a) + Bf\left(\frac{a+b}{2}\right) + Cf(b)$ has algebraic degree of exactness $\geqslant 2$. We get

$$\begin{cases} f(x)=1: & I(1)=\int_a^b 1dx=b-a=A\cdot 1+B\cdot 1+C\cdot 1 \\ f(x)=x-a: & I(x-a)=\int_a^b(x-a)dx=\frac{(b-a)^2}{2}=A\cdot 0+B\left(\frac{b-a}{2}\right)+C(b-a) \\ f(x)=(x-a)^2: & I\left((x-a)^2\right)=\int_a^b(x-a)^2dx=\frac{(b-a)^3}{3}=A\cdot 0+B\left(\frac{b-a}{2}\right)^2+C(b-a)^2 \end{cases}.$$

Solving this we get $A = \frac{1}{6}(b-a)$, $\quad B = \frac{4}{6}(b-a)$, $\quad C = \frac{1}{6}(b-a)$, i.e.

$$I(f)=\int_a^b f(x)dx \approx \frac{b-a}{6}\left[f(a)+4f\left(\frac{a+b}{2}\right)+f(b)\right]=\tilde{I}(f).$$

What is the algebraic degree of exactness? We know it's exact for $1, x, x^2$. Let's check for $f = x^3, x^4, \ldots$ until there is no equality.

$$f(x)=(x-a)^3: I\left(x^3\right)=\int_a^b(x-a)^3dx=\frac{(b-a)^4}{4}$$

$$\tilde{I}\left(x^3\right)=\left(\frac{b-a}{6}\right)\cdot 0+\frac{4(b-a)}{6}\underbrace{\left(\frac{b-a}{2}\right)^3}_{f((a+b)/2)}+\frac{b-a}{6}\underbrace{(b-a)^3}_{f(b)}=\frac{(b-a)^4}{4}.$$

Check that $I\left(x^4\right) \neq \tilde{I}\left(x^4\right)$. So the algebraic degree of exactness=3.

Error computation: check that we are in **Case B**: $\int_a^b q_3(x)dx = \int_a^b (x-a)(x-b)(x-(a+b)/2)dx = 0$. Then choosing $x_{n+1} = x_3 = \frac{a+b}{2}$, we get that $q_3(x)(x-x_3)$ does not change sign, and so by (3) we have

$$I(f) = \int_a^b f(x)dx = \underbrace{\int_a^b P_2(x)dx}_{I(P_2)} + \underbrace{\int_a^b f\left[a,b,\frac{a+b}{2},x\right]\psi(x)\,dx}_{E(I(f))=error} = I(P_2) + E(I(f))$$

$$= \underbrace{\frac{b-a}{6}\left[f(a) + 4f\left(\frac{a+b}{2}\right) + f(b)\right]}_{I(f)=\text{Simpson's Integration Rule}} + \underbrace{\int_a^b f\left[a,b,\frac{a+b}{2},x\right](x-a)(x-b)\left(x-\frac{a+b}{2}\right)^2 dx}_{E(I(f))=\text{ error}}$$

$$\begin{aligned} E(I(f)) &= \frac{f^{(4)}(\xi)}{24}\int_a^b (x-a)(x-b)\left(x-\frac{a+b}{2}\right)^2 dx \\ &= -\frac{f^{(4)}(\xi)(b-a)^5}{90\cdot 2^5}. \end{aligned}$$

Put $b-a=2h$ then $E(f) = -\frac{(2h)^5}{2^5\cdot 90}f^{(4)}(\xi) = -\frac{h^5}{90}f^{(4)}(\xi)$ so approximation order=5.

Example 5 (Corrected Trapezoidal Rule)**.** We can also use Hermite's interpolaton polynomial: $x_0 = x_1 = a$, $x_2 = x_3 = b$ so $n=3$ and

$$\begin{aligned} P_3(x) &= f[a] + f[a,a](x-a) + f[a,a,b](x-a)^2 + f[a,a,b,b](x-a)^2(x-b) \\ &= f(a) + f'(a)(x-a) + \frac{(f(b)-f(a))/(b-a) - f'(a)}{b-a}(x-a)^2 + \\ &\quad + \frac{f'(b) - 2(f(b)-f(a))/(b-a) + f'(a)}{(b-a)^2}(x-a)^2(x-b) \end{aligned}$$

$$\int_a^b f(x)dx = \int_a^b P_3(x)dx + \int_a^b f[a,a,b,b,x]q_4(x)dx$$

$$\int_a^b P_3(x)dx = \underbrace{\frac{b-a}{2}[f(a)+f(b)]}_{\text{Standard Trapezoidal rule}} + \frac{(b-a)^2}{12}\left[f'(a) - f'(b)\right] \quad \textbf{Corrected Trapezoidal Rule}$$

Compared with the Trapezoidal Rule, we have additional terms involving derivatives. We get an improvement in the error order: since $q_4(x) = (x-a)^2(x-b)^2 \geq 0$ we are in **Case A.** Put $h = b-a$, then by (2)

$$E(I(f)) = \frac{f^{(4)}(\xi)}{24}\underbrace{\int_a^b (x-a)^2(x-b)^2 dx}_{\frac{(b-a)^5}{30}}$$

$$= \frac{h^5}{720}f^{(4)}(\xi),$$

i.e. approximation order=5.

3 Composed/Composite Integration Rules

When $b-a$ is not small, the standard rules above will not result in a small error. Furthermore, as we mentioned in Interpolation Lecture Example 6, we may not even necessarily have $P_n \to f$ as $n \to \infty$ (for equispaced interpolation points). A standard method to reduce the error is to divide the interval $[a,b]$ into many sub-intervals (suppose with equispaced points)

$$\begin{aligned} & a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b, \quad x_i - x_{i-1} = h, \; i = 1, \ldots, n, \\ & h = \frac{b-a}{n} \end{aligned}$$

approximate $I_i(f) = \int_{x_i}^{x_{i+1}} f(x)\, dx$ and then sum up the approximations. The errors will also be added.

Example 6 (Composed Trapezoidal Rule)**.** We use the Trapezoidal Rule for each I_i:

$$\int_{x_i}^{x_{i+1}} f(x)\, dx = \underbrace{\frac{h}{2}(f(x_i) + f(x_{i+1}))}_{\text{Trapezoidal Integration Rule}} \underbrace{-\frac{f''(\xi_i)}{12} h^3}_{\text{error}}.$$

$$\int_a^b f(x) dx = \frac{h}{2} \sum_{i=0}^{n-1} \{f(x_i) + f(x_{i+1})\} - \frac{h^3}{12} \sum_{i=0}^{n-1} f''(\xi_i)$$

Note that each point $\{x_i\}_{i=1}^{n-1}$ appears twice in the sum while the endpoints appear once:

$$\frac{h}{2} \sum_{i=0}^{n-1} \{f(x_i) + f(x_{i+1})\} = \frac{h}{2} \left\{ f(x_0) + 2 \sum_{i=1}^{n-1} f(x_i) + f(x_n) \right\}. \quad \text{Composed Trapezoidal Rule}$$

If $f''(x)$ is continuous in $[a,b]$ we can use the intermediate value theorem and estimate the overall error by

$$\begin{aligned} -\sum_{i=0}^{n-1} \frac{h^3}{12} f^{(2)}(\xi_i) &= -\frac{nh^3}{12} \sum_{i=0}^{n-1} \frac{f''(\xi_i)}{n} = -\frac{(nh)h^2}{12} f''(\xi) \\ &= -\frac{(b-a)h^2}{12} f''(\xi). \end{aligned}$$

If $f''(x)$ is bounded but not necessarily continuous, $|f''(x)| \le M_2$ then

$$\begin{aligned} |E(I(f))| &\le M_2 \sum_{i=0}^{n-1} \frac{h^3}{12} = n \frac{M_2 h^3}{12} \\ &= \frac{M_2 h^2}{12} \end{aligned}$$

so in any case the approximation order=2. Note that this is 1 less than the standard Trapezoidal Rule.

Similarly we can get the composed Simpson's rule.

Example 7 (Composed Simpson's rule)**.** Now we need 3 points in each interval. So we divide $[a, b]$ into $2n$ intervals of equal length

$$a = \underbrace{x_0 < x_1 < x_2}_{\text{Interval for Simpson}} < \cdots < \underbrace{x_{2n-2} < x_{2n-1} < x_{2n}}_{\text{Interval for Simpson}} = b, \quad h = \frac{b-a}{2n}$$

and apply Simpson's rule in $[x_{2k-2}, x_{2k}]$ for $k = 1, 2, \ldots, n$:

$$\begin{aligned}
\int_a^b f(x)dx &= \sum_{k=1}^{n} \int_{x_{2k-2}}^{x_{2k}} f(x)dx = \sum_{k=1}^{n} \frac{2h}{6} \left[f\left(x_{2k-2}\right) + 4f\left(x_{2k-1}\right) + f\left(x_{2k}\right)\right] + \sum_{k=1}^{n} E\left(I_k\left(f\right)\right) \\
&= \frac{b-a}{6n} \sum_{k=1}^{n} \left[f\left(x_{2k-2}\right) + 4f\left(x_{2k-1}\right) + f\left(x_{2k}\right)\right] - \frac{h^5}{90} \sum_{k=1}^{n} f^{(4)}\left(c_k\right) \\
&= \frac{b-a}{6n} \left\{ f\left(x_0\right) + 2\sum_{k=1}^{n-1} f\left(x_{2k}\right) + 4\sum_{k=1}^{n} f\left(x_{2k-1}\right) + f\left(x_{2n}\right) \right\} - \frac{2nh}{180} h^4 \sum_{k=1}^{n} \frac{f^{(4)}\left(c_k\right)}{n} \\
&= \underbrace{\frac{h}{3} \left\{ f\left(x_0\right) + 2\sum_{k=1}^{n-1} f\left(x_{2k}\right) + 4\sum_{k=1}^{n} f\left(x_{2k-1}\right) + f\left(x_{2n}\right) \right\}}_{\text{Composed Simpson's Rule}} - \frac{b-a}{180} h^4 f^{(4)}\left(c\right).
\end{aligned}$$

4 Sensitivity to errors in integration

Consider Composed Trapezoidal Rule:

$$T\left(h\right) = \frac{h}{2} \left\{ f\left(x_0\right) + 2\sum_{i=1}^{n-1} f\left(x_i\right) + f\left(x_n\right) \right\}.$$

Suppose every $f\left(x_i\right)$ is known with some relative error $e_i = \frac{f(x_i) - \tilde{f}(x_i)}{f(x_i)}$ (this can be due to floating-point roundoff, measurement error etc.). Suppose $|e_i| \leq \varepsilon$. Then we compute

$$\begin{aligned}
\tilde{T}\left(h\right) &= \frac{h}{2} \left\{ \tilde{f}\left(x_0\right) + 2\sum_{i=1}^{n-1} \tilde{f}\left(x_i\right) + \tilde{f}\left(x_n\right) \right\} \\
\left|T\left(h\right) - \tilde{T}\left(h\right)\right| &= \frac{h}{2} \left\{ e_0 f\left(x_0\right) + 2\sum_{i=1}^{n-1} e_i f\left(x_i\right) + e_n f\left(x_n\right) \right\} \\
&\leqslant nh\varepsilon \max_{x\in[a,b]} \left|f\left(x\right)\right| \\
&= \left(b-a\right)\varepsilon M_0
\end{aligned}$$

So we get overall error of the same order as every individual error, therefore the integration (in general) is insensitive to errors in data, and we can take $h \ll 1$ (this is not true for numerical differentiation!)