Quadrature Methods for Numerical Integration

Tony Saad*
Institute for Clean and Secure Energy
University of Utah

April 11, 2011

1 The Need for Numerical Integration

Nuemrical integration aims at approximating definite integrals using numerical techniques. There are many situations where numerical integration is needed. For example, several well defined functions do not have an anti-derivative, i.e. their anti-derivative cannot be expressed in terms of primitive function. A popular example is the function e^{-x^2} whose anti-derivative does not exist. This function arises in a variety of applications such as those related to propability and statistics analyses. Furthermore, many applications in science and engineering are represented by integro-differential equations that require a special treatment for the integral terms (e.g. expansion, lineralization, closure ...). Therefore, numerical integration does not only provide a means for evaluating integrals numerically, but also grants us the ability to approximate special functions that are defined in terms of integrals.

Without loss of generality, there are two classes of problems where numerical integration is needed. In the first class, one wishes to evaluate the integral of a well defined function. In this case, the integrand can be evaluated a various points because and numerical integration techniques help define the optimum number of these points as well as their locations.

The second class of problems for applying numerical integration is found in differential equations the most common of which are those that express conservation principles. For example, the population balance equation, a well known partial differential equation encountered in process modeling and biological systems, exhibits source terms that are represented as integrals of the solution variable (e.g. the number density function).

The most common technique for numerical integration is called quadrature. The recipe for quadrature consists of three steps

1. Approximate the integrand by an interpolating polynomial using a specified number of points or nodes
2. Substitute the interpolating polynomial into the integral
3. Integrate

The resulting quadrature approximates the integral as a summation of the form

$$\int_a^b f(x)\mathrm{d}x = \sum_{i=1}^{n} w_i f(x_i). \tag{1}$$

*Postdoctoral Fellow. web:http://www.tsaad.net. email:saadtony@gmail.com.

Furthermore, if the nodes are selected in a specific manner using orthogonal polynomials, the accuracy of the quadrature formula is substantially improved as will be shown in subsequent sections.

I will begin our study of quadrature methods by reviewing the theory of interpolating polynomials. Then, I will introduce quadrature approximations for equally spaced nodes. This is followed by a discussion of the theory of orthogonal polynomials. Finally, I will show you how orthogonal polynomials can help in improving the degree of exactness of quadrature approximations.

2 Interpolating Polynomials

Interpolating polynomials are used to approximate a real valued function $f(x)$ on a real interval $[a,b]$ by a polynomial $p(x)$ that accurately represents the function over that interval.

2.1 Lagrange Interpolation

The most commonly used interpolating polynomials are known as Lagrange polynomials. Consider a real, continuous function $f(x)$ on a real interval $[a,b]$. Also, assume that the values of this function are known at a finite number of points $x_i \in [a,b]$, then, the Lagrange interpolating polynomial $p_n(x)$ of order n is the lowest degree polynomial such that $p_n(x_i) = f(x_i)$. Assuming that the interval is subdivided into $(n+1)$ points $(x_0 < x_1 < \cdots < x_n \in [a,b])$, then $p_n(x)$ is given by

$$p_n(x) = \sum_{i=0}^{n} f(x_i) L_i(x), \tag{2}$$

where $L_i(x)$ is the i-th Lagrange interpolating polynomial defined as

$$L_i(x) = \frac{\prod_{j=0, j\neq i}^{n} (x - x_j)}{\prod_{j=0, j\neq i}^{n} (x_i - x_j)}. \tag{3}$$

The first three Lagrange polynomials are given by

$$\begin{aligned}
L_0(x) &= \frac{x-x_1}{x_0-x_1}\frac{x-x_2}{x_0-x_2}\frac{x-x_3}{x_0-x_3}\cdots \\
L_1(x) &= \frac{x-x_0}{x_1-x_0}\frac{x-x_2}{x_1-x_2}\frac{x-x_3}{x_1-x_3}\cdots \\
L_2(x) &= \frac{x-x_0}{x_2-x_0}\frac{x-x_1}{x_2-x_1}\frac{x-x_3}{x_2-x_3}\cdots \\
&\vdots
\end{aligned} \tag{4}$$

2.2 Example

Find the interpolating polynomial for the following set of points

$$\begin{array}{cc}
x_0 = -1.5 & f(x_0) = -14.1014 \\
x_1 = -0.75 & f(x_1) = -0.931596 \\
x_2 = 0 & f(x_2) = 0 \\
x_3 = 0.75 & f(x_3) = 0.931596 \\
x_4 = 1.5 & f(x_4) = 14.1014.
\end{array} \tag{5}$$

The fundamental lagrangian polynomials are

$$L_0(x) = \frac{x-x_1}{x_0-x_1}\frac{x-x_2}{x_0-x_2}\frac{x-x_3}{x_0-x_3}\frac{x-x_4}{x_0-x_4} = \frac{1}{243}x(2x-3)(4x-3)(4x+3), \tag{6}$$

$$L_1(x) = \frac{x-x_0}{x_1-x_0}\frac{x-x_2}{x_1-x_2}\frac{x-x_3}{x_1-x_3}\frac{x-x_4}{x_1-x_4} = -\frac{8}{243}x(2x-3)(2x+3)(4x-3), \tag{7}$$

$$L_2(x) = \frac{x-x_0}{x_2-x_0}\frac{x-x_1}{x_2-x_1}\frac{x-x_3}{x_2-x_3}\frac{x-x_4}{x_2-x_4} = \frac{3}{243}(2x+3)(4x+3)(4x-3)(2x-3), \tag{8}$$

$$\vdots \tag{9}$$

2.3 Interpolation Error

The remainder or interpolation error for using the Lagrange interpolating polynomials is given by

$$R_n(x) = f(x) - P_n(x) = \frac{f^{n+1}(\xi)}{(n+1)!}\prod_{j=0}^{n+1}(x-x_j); \quad \xi \equiv \xi(x), \tag{10}$$

where $\xi \in (a,b)$.

3 Quadrature

Numerical integration is based on the idea of first approximating the integrand using an interpolating polynomial and then integrating the resulting polynomial. Assume that we wish to calculate the following integral

$$\int_a^b f(x)\mathrm{d}x. \tag{11}$$

Let

$$f(x) \approx p_{n-1}(x) = \sum_{i=1}^{n} f(x_i)L_i(x), \tag{12}$$

denote the Lagrange interpolating polynomial for $f(x)$. This is a polynomial of order $(n-1)$ given the n node approximation. Note that $L_i(x)$ is given by

$$L_i(x) = \frac{\prod_{j=1,j\neq i}^{n}(x-x_j)}{\prod_{j=1,j\neq i}^{n}(x_i-x_j)}. \tag{13}$$

We now substitute the interpolating polynomial into the integral

$$\int_a^b f(x)\mathrm{d}x \approx \int_a^b \sum_{i=1}^{n} f(x_i)L_i(x)\mathrm{d}x = \sum_{i=1}^{n} f(x_i)\int_a^b L_i(x)\mathrm{d}x = \sum_{i=1}^{n} w_i f(x_i), \tag{14}$$

where

$$w_i \equiv \int_a^b L_i(x)\mathrm{d}x. \tag{15}$$

This formula is known as a quadrature approximation for an integral. The points x_i are referred to as the abscissae or nodes while w_i are called the weights.

3.1 Degree of Exactness

A quadrature approximation is said to have degree of exactness m if it is exact when $f(x)$ is a polynomial of degree less than or equal to m, while it is not exact for a polynomial of order $m+1$. As a rule of thumb, any interpolatory quadrature formula that uses n distinct nodes has degree of exactness of at least $n-1$.

3.2 Gaussian Quadrature

Gaussian quadrature aims at improving the degree of exactness of the quadrature approximation by carefully selecting the abscissae of the quadrature formula. It also generalizes the concept of quadrature to integrals of the form

$$\int_a^b f(x)w(x)\mathrm{d}x, \tag{16}$$

where $w(x)$ is a weight function. A weight function $w(x)$ is a positive measurable function on a domain Ω such that

$$w(x) : \Omega \to \mathbb{R}^+. \tag{17}$$

It also has the following property

$$\int_a^b |x|^n w(x)\mathrm{d}x < \infty; \quad n = 0, 1, 2, \ldots \tag{18}$$

Using an $n-$point quadrature rule for Eq. (16), we have

$$\int_a^b f(x)\mathrm{d}x = \sum_{i=1}^n w_i f(x_i); \quad w_i \equiv \int_a^b L_i(x)w(x)\mathrm{d}x; \quad a \le x_1 < x_2 < \cdots < x_n \le b. \tag{19}$$

Regardless of how we choose the n abscissae, this quadrature approximation has degree of exactness at least equal to $(n-1)$. With Gaussian quadrature, one can achieve a degree of exactness of more than twice! Before we see how how this is possible, we'll have to go through some aspects of the theory of orthogonal polynomials.

4 Orthogonality

Two real function $f(x)$ and $g(x)$ are said to be orthogonal if their inner product is zero. The inner product of two functions, on an interval $[a,b]$, is defined by the following integral

$$\langle f, g \rangle \equiv \int_a^b f(x)g(x)\mathrm{d}x. \tag{20}$$

Then, $f(x)$ and $g(x)$are orthogonal if

$$\langle f, g \rangle = 0. \tag{21}$$

The above convolution is also known as an inner product.

4.1 Orthogonal Polynomials

The ideas of orthogonal functions can be used to construct a set of polynomials that can be used as a basis spanning a space of real functions. As a result, every function in that space can be written as a linear combination of the orthogonal basis. But we will not be concerned with group theory at this point, and we can proceed to developing orthogonal polynomials.

Consider a sequence of polynomials $p_k(x)$ such that

$$p_k(x) = \sum_{i=0}^{k} \alpha_{k,i} x^i; \quad \alpha_{k,k} = 1,\ k = 0, 1, 2, \cdots \tag{22}$$

For example,

$$p_0(x) = \alpha_{0,0} = 1, \tag{23}$$

$$p_1(x) = \alpha_{1,0} + \alpha_{1,1} x = x + \alpha_{1,0}, \tag{24}$$

$$p_2(x) = x^2 + \alpha_{2,1} x + \alpha_{2,0}. \tag{25}$$

By setting $\alpha_{k,k} = 1$, the polynomials are said to be monic, i.e. the coefficient of the term with highest order is one.

A sequence of polynomials $\mathscr{P} = \{p_m(x);\ m = 0, 1, \ldots, \infty\}$ is said to be orthogonal if

$$\begin{cases} \langle p_n, p_m \rangle = 0 & \forall n \neq m \\ \langle p_n, p_m \rangle \neq 0 & n = m \end{cases}, \tag{26}$$

or, in compact form

$$\langle p_n, p_m \rangle = \delta_{nm} M_n, \tag{27}$$

where δ_{nm} is the Kronecker delta and $M_n = \langle p_n, p_n \rangle$.

If, in addition, $\langle p_n, p_n \rangle = 1$, then the polynomials are said to be orthonormal. Therefore, an orthonormal set of polynomials is a normalized set of orthogonal polynomials. Orthonormal polynomials are defined using the following compact notation

$$\langle p_n, p_m \rangle = \delta_{nm}. \tag{28}$$

One can define a sequence of orthonormal polynomials $q_k(x)$ by normalizing the orthogonal ones as

$$q_k(x) = \frac{p_k(x)}{\sqrt{\langle p_k, p_k \rangle}}. \tag{29}$$

You can immediately verify that

$$\langle q_n, q_m \rangle = \langle \frac{p_n(x)}{\sqrt{\langle p_n, p_n \rangle}}, \frac{p_m(x)}{\sqrt{\langle p_m, p_m \rangle}} \rangle = \frac{1}{\sqrt{\langle p_n, p_n \rangle}\sqrt{\langle p_m, p_m \rangle}} \langle p_n, p_m \rangle = \delta_{nm}. \tag{30}$$

hence, $q_k(x)$ are orthonormal.

4.2 Conctructing Orthogonal Polynomials

On can construct a sequence of orthogonal polynomials using the following three-term recurrence relation (TTRR)

$$
\begin{aligned}
p_{-1}(x) &= 0, \\
p_0(x) &= 1, \\
p_{n+1}(x) &= (x-\alpha_n)p_n(x) - \beta_n p_{n-1}(x).
\end{aligned}
\tag{31}
$$

The coefficients can be calculated by using orthogonality. First, for α_n, we multiply Eq. (31) by p_n

$$p_n p_{n+1} = (x-\alpha_n)p_n p_n - \beta_n p_n p_{n-1}. \tag{32}$$

Next, we integrate over $[a,b]$

$$\int_a^b p_n p_{n+1}\mathrm{d}x = \int_a^b (x-\alpha_n)p_n p_n \mathrm{d}x - \int_a^b \beta_n p_n p_{n-1}\mathrm{d}x. \tag{33}$$

But, by virtue of orthogonality, we set

$$0 = \int_a^b (x-\alpha_n)p_n p_n \mathrm{d}x - 0, \tag{34}$$

or

$$0 = \int_a^b x p_n p_n \mathrm{d}x - \alpha_n \int_a^b p_n p_n \mathrm{d}x, \tag{35}$$

finally

$$\alpha_n = \frac{\int_a^b x p_n p_n \mathrm{d}x}{\int_a^b p_n p_n \mathrm{d}x} = \frac{\langle x p_n, p_n\rangle}{\langle p_n, p_n\rangle}. \tag{36}$$

For β_n, we multiply Eq. (31) by p_{n-1}

$$p_{n-1}p_{n+1} = (x-\alpha_n)p_{n-1}p_n - \beta_n p_{n-1}p_{n-1}. \tag{37}$$

Again, by integrating over $[a,b]$, we have

$$\int_a^b p_{n-1}p_{n+1}\mathrm{d}x = \int_a^b (x-\alpha_n)p_{n-1}p_n \mathrm{d}x - \int_a^b \beta_n p_{n-1}p_{n-1}\mathrm{d}x. \tag{38}$$

Using orthogonality, we write

$$0 = \int_a^b x p_{n-1}p_n \mathrm{d}x - \beta_n \int_a^b p_{n-1}p_{n-1}\mathrm{d}x, \tag{39}$$

finally

$$\beta_n = \frac{\int_a^b x p_{n-1}p_n \mathrm{d}x}{\int_a^b p_{n-1}p_{n-1}\mathrm{d}x} = \frac{\langle x p_{n-1}, p_n\rangle}{\langle p_{n-1}, p_{n-1}\rangle}. \tag{40}$$

At this point, it would be easier to find a simpler form for the term xp_{n-1} that appears in the numerator of Eq. (40). First, we observe that xp_{n-1} is a polynomial of order n. Also, because the polynomials are monic, we have

$$p_n - xp_{n-1} = \sum_{i=0}^{n-1} d_{n,i} p_i(x) \equiv q(x) \in \mathbb{P}_{n-1}. \tag{41}$$

In other wods, the difference $p_n - xp_{n-1}$ is a polynomial of order $(n-1)$ and can be written as a linear combination of all the lower order orthogonal polynomials. In fact, one can determine the coefficients of this linear combination very easily by using orthogonality. For example, for $m \leq n-1$, we form the following inner products

$$\langle p_n, p_m \rangle - \langle xp_{n-1}, p_m \rangle = \sum_{i=0}^{n-1} d_{n,i} \langle p_i, p_m \rangle, \tag{42}$$

or

$$0 - \langle xp_{n-1}, p_m \rangle = d_{n,m} \langle p_m, p_m \rangle, \tag{43}$$

so that

$$d_{n,m} = -\frac{\langle xp_{n-1}, p_m \rangle}{\langle p_m, p_m \rangle}. \tag{44}$$

At the outset, we can write the following

$$xp_{n-1} = p_n + q(x); \quad q(x) \in \mathbb{P}_{n-1}, \tag{45}$$

then, by taking the inner product, we recover

$$\langle xp_{n-1}, p_n \rangle = \langle p_n, p_n \rangle + \langle q, p_n \rangle = \langle p_n, p_n \rangle. \tag{46}$$

By substituting Eq. (46) into Eq. (31), the formula for calculating β_n is at hand

$$\beta_n = \frac{\langle p_n, p_n \rangle}{\langle p_{n-1}, p_{n-1} \rangle}. \tag{47}$$

4.3 Generalization

Orthogonal polynomials can also be defined with respect to a weight function $w(x)$. Two polynomials are said to be orthogonal with respect to a weight function $w(x)$ if

$$\int_a^b p_n(x) p_m(x) w(x) \mathrm{d}x = \delta_{nm} M_n = \begin{cases} 0 & n \neq m \\ M_n & n = m \end{cases}. \tag{48}$$

In a similar fashion, the orthogonal polynomials can be determined using the TTRR given in Eq. (31).

To calculate the coefficients α_n and β_n, we impose orthogonality with respect to $w(x)$. Starting with α_n, we multiply Eq. (31) by $p_n(x)$

$$p_n p_{n+1} w = (x - \alpha_n) p_n p_n w - \beta_n p_n p_{n-1} w. \tag{49}$$

Integrating over $[a, b]$, we have

$$\int_a^b p_n p_{n+1} w \mathrm{d}x = \int_a^b (x-\alpha_n) p_n p_n w \mathrm{d}x - \int_a^b \beta_n p_n p_{n-1} w \mathrm{d}x, \tag{50}$$

and, by virtue of orthogonality, we recover

$$0 = \int_a^b (x-\alpha_n) p_n p_n w \mathrm{d}x - 0, \tag{51}$$

or

$$0 = \int_a^b x p_n p_n w \mathrm{d}x - \alpha_n \int_a^b p_n p_n w \mathrm{d}x, \tag{52}$$

finally

$$\alpha_n = \frac{\int_a^b x p_n p_n w \mathrm{d}x}{\int_a^b p_n p_n w \mathrm{d}x} = \frac{\langle x w p_n, p_n \rangle}{\langle w p_n, p_n \rangle}. \tag{53}$$

For β_n, we multiply Eq. (31) by p_{n-1}

$$w p_{n-1} p_{n+1} = (x-\alpha_n) w p_{n-1} p_n - \beta_n w p_{n-1} p_{n-1}, \tag{54}$$

or

$$\int_a^b w p_{n-1} p_{n+1} \mathrm{d}x = \int_a^b w (x-\alpha_n) p_{n-1} p_n \mathrm{d}x - \int_a^b \beta_n w p_{n-1} p_{n-1} \mathrm{d}x, \tag{55}$$

then

$$0 = \int_a^b x w p_{n-1} p_n \mathrm{d}x - \beta_n \int_a^b w p_{n-1} p_{n-1} \mathrm{d}x, \tag{56}$$

finally

$$\beta_n = \frac{\int_a^b x w p_{n-1} p_n \mathrm{d}x}{\int_a^b w p_{n-1} p_{n-1} \mathrm{d}x} = \frac{\langle x w p_{n-1}, p_n \rangle}{\langle w p_{n-1}, p_{n-1} \rangle}. \tag{57}$$

As we did for the non-weighted case, we can simplify the numericator for β_n by writing

$$x w(x) p_{n-1}(x) = w(x) x p_{n-1}(x) = w(x)[p_n + q]; \quad q \in \mathbb{P}_{n-1}. \tag{58}$$

We now form the inner product

$$\langle x w p_{n-1}, p_n \rangle = \langle w p_n, p_n \rangle + \langle w p_{n-1}, p_n \rangle = \langle w p_n, p_n \rangle. \tag{59}$$

Upon substitution into Eq. (57), we recover the formula for calculating β_n as

$$\beta_n = \frac{\langle w p_n, p_n \rangle}{\langle w p_{n-1}, p_{n-1} \rangle}. \tag{60}$$

Orthogonal functions have many other properties that are outside the scope of this review. I'll get back to those at a later occasion, but for now, we have enough information to go ahead and discuss how orthogonal polynomials can be used to improve the degree of exactness of the quadrature approximation.

5 Gaussian Quadrature

Back to Gaussian quadrature, we said that it aims at improving the degree of exactness of the quadrature approximation by carefully selecting the nodes of the quadrature formula. For an $n-$point quadrature, if the abscissae are selected such that they coincide with the roots of the corresponding orthogonal polynomial $p_{n-1}(x)$, then the quadrature approximation has a degree of exactness of $(2n-1)$. Let us see how this is possible.

Suppose that $f(x)$ is a polynomial of degree $m \leq 2n-1$. Then, one can write

$$f(x) = p_n(x)q(x) + r(x), \tag{61}$$

where q and r are polynomials of degree $\leq n-1$. By virtue of orthogonality, we have

$$\int_a^b f(x)w(x)\mathrm{d}x = \int_a^b p_n(x)q(x)w(x)\mathrm{d}x + \int_a^b r(x)w(x)\mathrm{d}x = \int_a^b r(x)w(x)\mathrm{d}x. \tag{62}$$

Now the quadrature formula is

$$\int_a^b f(x)w(x)\mathrm{d}x = \sum_{i=1}^{n} w_i f(x_i). \tag{63}$$

Now, suppose that the nodes are selected such that they coincide with the roots of $p_n(x)$, i.e. $p_n(x_i) = 0, \quad i = 1,2,\ldots,n$. Then,

$$f(x_i) = p_n(x_i)q(x_i) + r(x_i) = 0 + r(x_i) \tag{64}$$

and our interpolatory rule becomes

$$\int_a^b f(x)w(x)\mathrm{d}x = \sum_{i=1}^{n} w_i f(x_i) = \sum_{i=1}^{n} w_i r(x_i) \tag{65}$$

and thus the approximation is exact because $r(x)$ is a polynomial of degree $\leq n-1$ (remember, an n-point quadrature rule is exact for a polynomial of degree $(n-1)$).

As you can see, by using only n nodes, and by specifically choosing those nodes as the roots of the $n-$th order orthogonal polynomial, then the quadrature approximation has a degree of exactness of $(2n-1)$.