Orthogonal Polynomials, Quadratures & Sparse-Grid Methods for Probability Integrals

Dr. Abebe Geletu
May, 2010

Technische Universität Ilmenau,
Institut für Automatisierungs- und Systemtechnik
Fachgebiet Simulation und Optimale Prozesse
Topics

- Interpolatory Quadrature Rules
- Orthogonal Polynomials
- Embedded Quadrature Rules
- Quadrature Rules and Probability Integrals
- Full-Grid Tensor Product Multidimensional Integration
- Sparse-Grid Tensor Product Multidimensional Integration

**Part - I:** Quadrature Rules;

**Part - II:** Sparse-Grid Integration Methods
1. Interpolatory Quadrature Rules

For an integrable function $f : \mathbb{R} \to \mathbb{R}$ to compute the integral

$$I[f] = \int_{a}^{b} f(x)\phi(x)dx,$$

where

- $[a, b] \subset \mathbb{R}$ is a bounded or unbounded interval;
- $\phi : \mathbb{R} \to \mathbb{R}$ is a nonnegative weight function.

Sometimes the weight function need to be adjusted to conform to the properties of the integrand $f$, e.g. oscillatory behaviors, etc.; and the interval of integration $\Omega_1$. 


1.2. Some integrals with standard weight functions and intervals of integration

(a) \[ l_1[f] = \int_{-1}^{1} f(x) \, dx \Rightarrow \phi(x) = 1; \Omega_1 = [-1, 1] \]

(b) \[ l_2[f] = \int_{-1}^{1} f(x)(1 - x)^\alpha(1 + x)^\beta \, dx \]

\[ \Rightarrow \phi(x) = (1 - x)^\alpha(1 + x)^\beta, \alpha, \beta > -1; \]
\[ \Omega_1 = [-1, 1] \]

(c) \[ l_3[f] = \int_{-1}^{1} f(x) \left(1 - x^2\right)^{-\frac{1}{2}} \, dx \]

\[ \Rightarrow \phi(x) = \left(1 - x^2\right)^{-\frac{1}{2}}; \Omega_1 = [-1, 1] \]
(d) \[ I_4[f] = \int_0^{+\infty} f(x)x^\alpha e^{-x} \, dx \]

\[ \Rightarrow \phi(x) = x^\alpha e^{-x}, \alpha > -1; \Omega_1 = [0, +\infty) \]

(e) \[ I_5[f] = \int_{-\infty}^{+\infty} f(x)e^{-x^2} \, dx \]

\[ \Rightarrow \phi(x) = e^{-x^2}; \Omega_1 = [0, +\infty) \]

**Note:**
- transform other types of integrals to the standard ones.

**Example:**

\[ \int_{-1}^{1} \sin(x)e^{-x^2} \, dx \rightarrow \int_{-1}^{1} \left( \sin(x)e^{-x^2} \sqrt{1 - x^2} \right) \left( 1 - x^2 \right)^{-\frac{1}{2}} \, dx \]

\[ = f(x) \]
1.3. Interpolatory Quadrature Formulas

To construct an N-point quadrature formula (rule):

\[ Q_N^1[f] := \sum_{k=1}^{N} w_k f(x_k), \]

where the integration nodes \( \mathcal{X} = \{x_1, x_2, \ldots, x_N\} \) and weights \( \{w_1, w_2, \ldots w_N\} \) are constructed based on \( \Omega_1 = [a, b] \) and the weight function \( \phi \).

**Question:**

How to determine the \( 2N \) unknowns \( x_1, \ldots, x_N \) and \( w_1, \ldots, w_N \)?

**Requirement:** the weights \( w_1, w_2, \ldots, w_N \) should be non-negative to avoid numerical cancelations.
1.4. Efficient interpolatory quadrature rules

1. Gauss quadrature rules and their extensions
2. Curtis-Clenshaw quadrature rules
1.5. Polynomial Exactness

One measure of quality for a quadrature rule is its **polynomial exactness**.

- The $N$-point quadrature $Q^1_N[\cdot]$ is said to be exact for a polynomial $p_m$ of degree $m$ if

  $$I[p_m] = \int_a^b p_m(x)\phi(x)\,dx = Q^1_N[p_m] = \sum_{k=1}^N w_k p_m(x_k)$$

- The **degree of exactness** or **degree of accuracy** $d$ of an $N$-point quadrature rule is equal to the maximum degree polynomial for which $Q^1_N$ is exact.
1.6. Gauss Quadrature Rules and Orthogonal Polynomials

An N-point Gauss quadrature rule is constructed to achieve the largest possible polynomial exactness.

**Theorem (see Davis & Rabinowitz [1]):**

An N-point quadrature formula $Q^1_N$ has degree of exactness $d = 2N - 1$ if the integration nodes $x_1, x_2, \ldots, x_N$ are zeros of the $N$-th degree orthogonal polynomial $p_N(x)$ w.r.t. $\phi$ and $\Omega_1 = [a, b]$.

Hence, there is a polynomial $p_m$ with degree $m > 2N - 1$, such that $I[p_m] \neq Q^1_N[p_m]$.

**Example:**

a) A 3-point Gauss quadrature rule $Q^1_3$ has a polynomial exactness $d = 2 \times 3 - 1 = 5$
Orthogonal Polynomials

Two polynomials \( p_n \) and \( p_m \) are orthogonal w.r.t. \( \phi \) and \( \Omega_1 = [a, b] \) if

\[
< p_n, p_m > := \int_a^b p_n(x)p_m(x)\phi(x)dx = 0.
\]

In Theorem 1, the \( N \)-th degree orthogonal polynomial \( p_N \) is orthogonal to all polynomials \( p_m \) degree \( m \leq N - 1 \).

Examples:

(a) \( p_0 = 1, \ p_1(x) = x \) are orthogonal w.r.t. \( \phi(x) = 1 \) and \( \Omega_1 = [-1, 1] \).

(b) \( H_1 = 2x, \ H_2(x) = 4x^2 - 2 \) are orthogonal w.r.t. \( \phi(x) = e^{-x^2} \) and \( \Omega_1 = (-\infty, +\infty) \).
Characterization of Orthogonal Polynomials

Theorem (Recurrence relationship, see Gatushi [2])

Suppose that $p_0 = 1, p_1, \ldots$ are orthogonal polynomials with $\text{deg}(p_n) = n$ and leading coefficient equal to 1. For every integer $n$

$$p_{n+1}(x) = (x - a_n)p_n(x) - b_n p_{n-1}(x), \quad n = 1, 2 \ldots \quad (*)$$

where

$$a_n = \frac{\int_a^b x p_n^2(x) \phi(x) \, dx}{\int_a^b p_n^2(x) \phi(x) \, dx}, \quad n = 1, \ldots ;$$

$$b_n = \frac{\int_a^b x p_n^2(x) \phi(x) \, dx}{\int_a^b p_{n-1}^2(x) \phi(x) \, dx}, \quad n = 1, \ldots$$
Characterization of Orthogonal Polynomials

Theorem (Continued...Recurrence relation)

and

\[ p_1(x) = (x - a_0)p_0(x) = x - a_0, \]
\[ a_0 = \frac{\int_a^b xp_0^2(x)\phi(x)dx}{\int_a^b p_0^2(x)\phi(x)dx} = \frac{\int_a^b x\phi(x)dx}{\int_a^b \phi(x)dx}. \]

Remark:

- Given a nonnegative weight function \( \phi \) and a set \( \Omega \), you can construct your own set of orthogonal polynomials if you can efficiently compute the recurrence coefficients \( a_0, a_n, b_n, n = 1, 2, \ldots \).
1.7. Algorithms to determine the recurrence coefficients

- Algorithms for the computation of the recurrence coefficients are commonly known as Steleje's Procedures.
- Currently an efficient and stable algorithm for the computation of the coefficients $a_0, a_n, b_n, n = 1, 2, \ldots$ is given by Gander & Karp [5].

**Software:**
- FORTRAN: ORTHOPOL by Gatushi [4]
- C++ implementation of ORTHOPOL Fernandes & Atchley [4].
- Various Matlab codes by Gatushi [1]
Some known sets of orthogonal polynomials

- Jacobi polynomials orthogonal w.r.t.
  \[ \phi(x) = (1 - x)^\alpha (1 + x)^\beta, \quad \alpha, \beta > -1 \text{ and } \Omega_1 = [-1, 1]: \]
  \[
p^{(\alpha,\beta)}_n(x) = \frac{1}{2^n} \sum_{k=0}^{n} \binom{n + \alpha}{k} \binom{n + \beta}{n - k} (x - 1)^{n-k}(1 + x)^k, \]
  \[ n = 0, 1, 2, \ldots \]

- Lagendre polynomials are the Jacobi polynomials for
  \[ \alpha = \beta = 0; \text{ i.e. } \phi(x) = 1: \]
  \[
  L_n(x) = \frac{1}{2^n} \sum_{k=0}^{n} \binom{n}{k} \binom{n}{n - k} (x - 1)^{n-k}(1 + x)^k,
  \]
  \[ n = 0, 1, 2, \ldots \]
Hermite Polynomials w.r.t. $\phi = e^{-x^2}$ and $\Omega_1 = (-\infty, +\infty)$

$$H_n(x) = n! \sum_{k=1}^{\lfloor n/2 \rfloor} \frac{(-1)^k}{k!(n-2k)!} (2x)^{n-2k}, \quad n = 0, 1, 2, \ldots$$

The first 6 terms of Hermite polynomials:

$$H_0(x) = 1, \quad H_1(x) = 2x, \quad H_2(x) = 4x^2 - 2$$
$$H_3(x) = 8x^3 - 12x, \quad H_4(x) = 16x^4 - 48x^2 + 12,$$
$$H_5(x) = 32x^5 - 160x^3 + 120x$$
$$H_6(x) = 64x^6 - 480x^4 + 720x^2 - 120, etc.$$
Chebychev Polynomials orthogonal w.r.t. 
\( \phi(x) = (1 - x^2)^{-\frac{1}{2}} \) and \( \Omega_1 = [-1, 1] \)

\[
C_0(x) = \frac{1}{\sqrt{\pi}} T_0(x), \quad C_n(x) = \sqrt{\frac{2}{\pi}} T_n(x), \quad n = 1, 2, \ldots
\]

where \( T_n(x) = \cos(n \arccos(x)) \), known as Chebychev polynomials of first kind. Few terms are given by

\[
T_0(x) = 1, \quad T_1(x) = x, \quad T_2(x) = 2x^2 - 1, \\
T_3(x) = 4x^3 - 3x, \quad T_4(x) = 8x^4 - 8x^2 + 1 \\
T_5(x) = 16x^5 - 20x^3 + 5x, \quad T_6(x) = 32x^6 - 48x^4 + 18x^2 - 1
\]

Corresponding to each set of orthogonal polynomials we have the quadrature rules: Gauss-Jacobi, Gauss-Legendre, Gauss-Hermite, Gauss-Chebychev, etc.
1.8. Computation of quadrature nodes and weights

**Theorem (see Gatushi [2, 3])**

The quadrature nodes $x_1, \ldots, x_N$ and weights $w_1, w_2, \ldots, w_N$ can be obtained from the spectral factorization

$$J_n = V^\top \Lambda V; \quad \Lambda = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_N), \quad VV^\top = I_N;$$

of the symmetric matrix tridiagonal **Jacobi matrix**

$$J_n = \begin{bmatrix}
a_0 & \sqrt{b_1} & & \\
\sqrt{b_1} & a_1 & \sqrt{b_2} & \\
& \sqrt{b_2} & \ddots & \ddots \\
& & \ddots & a_{N-2} & \sqrt{b_{N-1}} \\
& & & \sqrt{b_{N-1}} & a_{N-1}
\end{bmatrix},$$
Theorem (continued...)

where \(a_k, b_k, k = 1, 2, \ldots, N - 1\) are known coefficients from the recurrence relation (★). In particular

\[
x_k = \lambda_k, \quad k = 1, 2, \ldots, N;
\]

\[
w_k = \left(e_1^\top V e_k\right)^2, \quad k = 1, 2, \ldots, N.
\]

- Observe that: all the weights \(w_k\) obtained above are nonnegative.

**Exercise:**
Write a Matlab code not longer than 10 lines to compute \(x_k'\)s and \(w_k'\)s.
1.9. Gauss Quadrature Rules with Preassigned nodes

- In some application integration nodes need to be prefixed (pre-given), e.g. due to boundary conditions, constraints, etc.

- Such nodes are usually needed in collocation methods in the solution of ODEs and PDEs.

**Objective:** To construct an \(N\)–point quadrature rule \(Q^1_N\) with \(x_1, \ldots, x_m, m \leq N\), are preassigned (fixed) nodes in \([a, b]\):
  - determine the remaining \(N - m\) nodes;
  - corresponding weights \(w_1, w_2, \ldots, w_N\)

so that \(Q^1_N\) has the maximum possible degree of exactness \(d\).
Classical Gauss quadrature rules with preassigned nodes

In a closed bounded interval \([a, b]\):

1. **Gauss-Radau rule:**
   - One end point of \([a, b]\) is prefixed to be a node; i.e. either \(x_1 = a\) or \(x_N = b\); the rest of the nodes \(x_2, x_3, \ldots, x_N \in (a, b)\).
   - Degree of polynomial exactness \(d = 2N - 2\).

2. **Gauss-Lobatto rule:**
   - Both end points of the interval \([a, b]\) are prefixed to be nodes; i.e. \(x_1 = a\) and \(x_N = b\). There rest \(x_2, \ldots, x_N \in (a, n)\).
   - Degree of polynomial exactness \(d = 2N - 3\).

In general, prefixing nodes reduces the degree of polynomial exactness by the number of integration nodes prefixed.
Recently, Bultheel et. al. [2] give Gauss quadrature rules with prefixed nodes and positive weights in any interval $\Omega_1 \subset \mathbb{R}$ bounded or unbounded.

**Software:**

1.10 How good are Gauss Quadrature Rules?

Disadvantage:

- Different Gauss quadrature rules $Q_{N_1}^1$ and $Q_{N_2}^1$ have only a few common nodes, for $N_1 < N_2$; i.e. the set of nodes

  $\mathcal{X}_1 = \{x_1, x_2, \ldots, x_{N_1}\} \subset [a, b]$ and

  $\mathcal{X}_2 = \{z_1, z_2, \ldots, z_{N_2}\} \subset [a, b]$.

  $\Rightarrow$ it is difficult to estimate the convergence of the limit:

  $$\lim_{N \to \infty} \left| I[f] - Q_N^1[f] \right|.$$ 

Note that: had it been $\mathcal{X}_1 \subset \mathcal{X}_2$, we could have obtained

  $$\left| I[f] - Q_{N_2}^1[f] \right| \leq \left| I[f] - Q_{N_1}^1[f] \right|.$$ 

Solution: Either extended Gauss quadrature rules so that the nodes for a lower accuracy can be reused to construct quadrature rules of higher accuracy; or construct embedded quadrature rules from the start.
1.11. Embedded quadrature rules - Extensions of Gauss quadrature rules

A) Kronord’s extension [6]: Given Gauss quadrature nodes \( x_1, x_2, \ldots, x_N \), between every two nodes add one new node:

- \( z_1 \in (a, x_1), z_2 \in (x_1, x_2), \ldots, z_N \in (x_{N-1}, x_N) \), \( z_{n+1} \in (x_N, b) \) so that the \( N + N + 1 = 2N + 1 \) points \( z_1, x_1, z_2, x_2, \ldots, x_{N-1}, z_N, x_N, z_{N+1} \) are the nodes of the new quadrature rules; and

- and the new weights \( w_1, w_2, \ldots, w_{2N+1} \) are all nonnegative.

\[ d = \begin{cases} 
3N + 1, & \text{if } N \text{ is even} \\
3N + 2, & \text{if } N \text{ is odd}
\end{cases} \]

\[ \mathcal{X}_N \subset \mathcal{X}_{2N+1}. \]
B) Patterson’s extension [2]:
To an existing $N$ Gauss quadrature nodes $x_1, x_2, \ldots, x_N$, add $m$ new integration nodes $z_1, z_2, \ldots, z_m$ so the new quadrature rule has the maximum degree of accuracy

$$d = 2(m + N) - N - 1 = N + 2m - 1$$

and the weights $w_1, w_2, \ldots, w_{N+m}$ are nonnegative.
Hence, $\mathcal{X}_N \subset \mathcal{X}_{N+m}$.

Note:
- In general, for a given weight function $\phi$ and integration domain $[a, b]$, the construction of convenient embedded Gauss-quadrature rules is not a trivial task.(see, [3, 3, 5])

Software:
Embedded quadrature rules - the Curtis-Clenshaw quadrature rule

For integrals on $[-1, 1]$ (in general on a closed and bounded interval $[a, b]$) the set of nodes

$$
\mathcal{X}_N = \left\{ x_k \mid x_k = \cos \left( \frac{(k - 1)\pi}{N - 1} \right), k = 1, 2, \ldots, N \right\}
$$

and the weights are computed from the orthogonal Chebychev polynomials of first type $T_n(x) = \cos(n \arccos(x))$ using Theorem 4.

Properties:
- degree accuracy $N - 1$;
- $\mathcal{X}_N \subset \mathcal{X}_{2N-1}$;
- nodes and weight are very simple to compute; (see Trefethen [3], Waldvogel [4] for Matlab codes);
- despite lower polynomial exactness, comparable efficiency with Gauss quadrature rules (Trefethen [3]);
- extensive applications in spectral and pseudo-spectral...
1.12. Quadrature Rules and Probability Integrals

For a well behaved nonnegative function $\rho(x)$, if

$$\int_{-\infty}^{\infty} \rho(x)dx = 1,$$

then the expression $\mu(x) = \rho(x)dx$ defines a probability measure. And $\rho$ is termed a probability density function of $x$. In particular, for a random set $A$

$$Pr\{x \in A\} = \mu(x \in A) = \int_{-\infty}^{\infty} 1_A(x)\rho(x)dx,$$

where

$$1_A(x) = \begin{cases} 1, & \text{if } x \in A \\ 0, & \text{if } x \notin A. \end{cases}$$

For function $f$ of the random variable $x$, its expected value is

$$E[f] = \int_{-\infty}^{\infty} f(x)\rho(x)dx.$$
Standard probability density functions

Given a nonnegative weight function $\phi(x)$ with the property

$$\phi(x) = \begin{cases} > 0, & \text{for } x \in [a, b] \\ 0, & \text{if } x \notin [a, b]. \end{cases}$$

such that $\phi_0 = \int_a^b \phi(x) dx$ set $\rho(x) := \frac{\phi(x)}{\phi_0}$, then the expression $\mu(x) = \rho(x) dx$ defines a probability measure.

Example:

- For $\phi(x) = 1$ on $[a, b]$:

  $$\phi_0 = b - a \Rightarrow \rho(x) = \frac{1}{\phi_0} = \frac{1}{b - a} \rightarrow \text{the uniform density function.}$$
For $\phi(x) = e^{-x^2}$ on $(-\infty, +\infty)$:

$$\phi_0 = \int_{-\infty}^{+\infty} e^{-x^2} = \sqrt{\pi} \Rightarrow \rho(x) = \frac{\phi(x)}{\phi_0} = \frac{e^{-x^2}}{\sqrt{\pi}}.$$

Hence,

$$F(x) = \int_{-\infty}^{x} \frac{e^{-z^2}}{\sqrt{\pi}} dz = \int_{-\infty}^{x} \frac{e^{-\frac{1}{2}(\sqrt{2}z)^2}}{\sqrt{2}\sqrt{\pi}} \sqrt{2}dz$$

Setting $u = \sqrt{2}z$ we have $du = \sqrt{2}dz$ and

$$F(x) = \int_{-\infty}^{x} \frac{e^{-z^2}}{\sqrt{\pi}} dz = \int_{-\infty}^{x} \frac{e^{-\frac{1}{2}u^2}}{\sqrt{2\pi}} du.$$ 

$\Rightarrow F$ defines the standard normal distribution.
Literature


