Modern Methods for high-dimensional quadrature

WS 2013/14

A. Griewank, H. Leövey

VL:
Die. 11-13 RUD 25, 1.013
Mi. 9-11 RUD 25, 1.013

UE:
Die. 13-15 RUD 25, 1.013

http://www2.mathematik.hu-berlin.de/~gaggle/W1314/MQI/UEBUNG/quadrature.pdf

I Review of 1-dimensional Integration

- Polynomial interpolation
  - Newton and Lagrange
  - Error of interpolation
  - Hermite interpolation
- Integration based on interpolation:
  - Basic rules
  - Coefficient determination
  - Error analysis
  - Change of intervals
- Orthogonal Polynomials and Gaussian Integration:
  - Approximation theory and orthogonal systems
  - Practical Construction of orthogonal polynomials: 3-term recursion
  - Basic Idea and differences with numerical integration based on interpolation
  - Golub/Eigenvalue formulation of 3-term recursion
  - Basic theorems of Gaussian integration
- Romberg and adaptive integration (if time permits):
  - Romberg algorithm: Basic description and convergence
  - Adaptive quadrature: Basic description of subdivision schemes, quadrature selection and stop-criterias
- General error theory from Peano (truncation error):
  - Peano’s theorem for Newton-Cotes and Gauss rules

II Monte Carlo and Quasi Monte Carlo Methods

- Product rules, curse of dimensionality
- Monte Carlo (MC) methods:
  - Classical/plain Monte Carlo integration
  - Pseudo-random number generation
- Inverse-transform method and acceptance rejection
- Importance sampling
- Variance reduction techniques

- Discrepancy and uniform distributed sequences:
  - Basic results and definitions
  - Relation to high dimensional integration (Koksma-Hlawka inequality)

- Quasi-Monte Carlo (QMC) Methods:
  - Elemental intervals and low discrepancy sequences
  - (t,m,d)-nets and (t,d)-sequences. Digital nets and sequences
  - Discrepancy upper-bounds for (t,m,d)-nets and (t,d)-sequences
  - Basic constructions: Halton, Sobol.
  - Lattice Rules
  - Integration error for lattice rules for periodic integrands

- Applications of QMC to Finance, particle physics, stochastic programming
  - Brownian motion paths generation methods
  - The quartic oscillator
  - Two-stage stochastic programs

- Weighted Reproducing Kernel Hilbert Spaces (WRKHS) (if time permits)
  - Setting and tractability
  - Results for Lattice rules and digital nets

### III Variations and Extensions of QMC

- Randomization of QMC (if time permits)
  - Randomly shifted lattice rules and errors in (WRKHS) of non-periodic nature
  - Randomly digitally shifted (t,m,d)-nets and (t,d)-sequences.
  - Random Scrambling of (t,m,d)-nets and (t,d)-sequences
  - Practical error estimation

- Sparse grids (if time permits)
  - Basic constructions
  - Error bounds for integration

- Analysis of Variance (ANOVA):
  - Definition of functional ANOVA decomposition
  - Properties
  - Effective dimensions
1 Review of 1-D integration

1.1 Lagrangian Interpolation

Proposition 1.1 Suppose \((x_i, y_i)\) with \(a \leq x_0 < x_1 \cdots < x_n \leq b\) and \(y_i = f(x_i)\) with \(f \in C^{n,1}([a,b])\). Then the unique interpolant

\[
P_n(x) = \sum_{i=0}^{n} f(x_i)\ell_i(x)
\]

with \(\ell_i(x) = \prod_{i \neq j=0}^{n} \frac{x-x_j}{x_i-x_j}\)

satisfies

\[
|f(x) - P_n(x)| \leq \frac{L}{(n+1)!} \left| \prod_{i=0}^{n} (x-x_i) \right|_q(x),
\]

where \(L\) is Lipschitz constant of \(f^{(n)}\), possibly

\[
L = \sup_{a \leq x \leq b} |f^{(n+1)}(x)|.
\]

Proof: If \(x = x_i\) for some \(0 \leq i \leq n\), then equality holds. Suppose now \(x \neq x_i\) for all \(0 \leq i \leq n\). Then \(q_n(x) \neq 0\) and we can define the function \(\phi(t) = f(t) - P_n(t) - \lambda x q_n(t)\), with \(\lambda_x := \frac{f(x) - P_n(x)}{q_n(x)}\). Since \(\phi \in C^{n,1}([a,b])\), and \(\phi\) vanishes at \(n + 2\) distinct points \(x, x_0, x_1, \ldots, x_n\), by Rolle’s theorem we have that \(\phi’\) has at least \(n+1\) distinct zeros in \([a,b]\). By following this argument \(n\) times we obtain that \(\phi^n\) has at least two distinct zeros, say \(\xi_1, \xi_2, \) in \([a,b]\). But then we have that

\[
0 = \phi^n(\xi_1) - \phi^n(\xi_2) = f^n(\xi_1) - f^n(\xi_2) + P^n(\xi_1) - P^n(\xi_2) + \lambda_x(q^n(\xi_2) - q^n(\xi_1))
\]

Since the \(\text{deg}(P_n) = n\), then \(P^n(\xi_2) - P^n(\xi_1) = 0\). Since \(\text{deg}(q_n) = n + 1\) and \(q_n\) monic, then \(q^n(\xi) = (n+1)!t + a\), for some \(a \in \mathbb{R}\). Therefore it follows

\[
0 = f^n(\xi_1) - f^n(\xi_2) + \lambda_x(n+1)!(\xi_2 - \xi_1)
\]

Due to the Lipschitz condition on \(f^n\), we have

\[
|f(x) - P_n(x)| = |\lambda x q_n(x)| = \left| \frac{f^n(\xi_2) - f^n(\xi_1)}{(n+1)!(\xi_2 - \xi_1)} q_n(x) \right| \leq \frac{L}{(n+1)!} |q_n(x)|.
\]

Remark: Smoothness = Differentiability Requirement usually quite high for classical higher order methods, but much lower for many modern applications and suitable methods.

Consequence: Bound of quadrature at Chebyshev points

\[
\|f - P_n\|_\infty = \sup_{a \leq x \leq b} |f(x) - P_n(x)| \leq |q_n|_\infty \frac{L}{(n+1)!},
\]

where \(q_n(x) = (x-x_0)(x-x_1), \ldots, (x-x_n)\) is arbitrary except for being monic, i.e., having the leading term \(x^n\).
Lemma 1.2 For $[a, b] = [-1, 1]$, $\|q_n\|_\infty$ is minimized by the Chebyshev polynomial

\[
q_n(x) = \frac{1}{2^n} T_{n+1}(x) \quad \text{where}
\]

\[
T_n(x) = \cos(n \arccos(x)) \in [-1, 1].
\]

\[\text{Proof:}\] See any numerics text book.

Substitution in above formula yields

\[
\|f - P_n\|_\infty \leq \frac{L}{(n + 1)!2^n},
\]

when interpolation is carried out at Chebyshev points, i.e., $n + 1$ roots of $T_{n+1}$.

1.2 Simple Quadratures based on Interpolation

Suppose we can pick $A_i$ for $i = 0, \ldots, n$ such that

\[
Q_n(f) := \sum_{i=0}^{n} A_i f(x_i) \approx \int_{a}^{b} f(x)dx
\]

holds exactly for all polynomials of degree $\leq n$. Then, if $P$ interpolates $f$ on the points $(x_i)_{i=0,\ldots,n}$, we have

\[
\left| \int_{a}^{b} f(x)dx - Q_n(f) \right| \leq \left| \int_{a}^{b} f(x)dx - \int_{a}^{b} P_n(x) \right|
\]

\[
\leq \frac{L}{(n+1)!} \int_{a}^{b} |q_n(x)|dx \leq \frac{L(b-a)}{(n+1)!} \|q_n(x)\|_\infty.
\]

When $[a, b] = [-1, 1]$, the last bound is minimized according to Lemma 1.2 by

\[
\frac{L}{(n + 1)!2^n} = \frac{L}{(n+1)! \cdot 2^{n-1}}
\]

Lemma 1.3 When $[a, b] = [-1, 1]$, $\int_{-1}^{1} |q_n(x)|dx$ is minimal by the Chebyshev polynomial of the second kind

\[
q_n(x) = \frac{1}{2^{n+1}} U_{n+1}(x) \quad \text{with} \quad U_{n+1} = \frac{\sin((n+2) \arccos(x))}{\sqrt{1 - x^2}}
\]

for which $\int_{-1}^{1} |q_n(x)|dx = \frac{1}{2^{n+1}} \int_{1}^{1} |U_{n+1}|dx = \frac{1}{2^n}$.

\[\text{Consequence}\] Estimate for quadrature points $(x_i)_{i=0,\ldots,n}$, taken as the roots of the Chebyshev polynomial of second kind

\[
\left| \int_{-1}^{1} f(x)dx - Q_n(f) \right| \leq \frac{L}{(n + 1)!2^n}.
\]
Hence, all we gained is a factor of 2.

**Newton-Cotes formulas for uniform grid**

\[ x_i = -1 + 2i/n \quad \text{for } i = 0, \ldots, n \]

\[
\sup_{-1 \leq x \leq 1} \left| \prod_{i=0}^{n} (x - x_i) \right| \leq \prod_{i=1}^{n} (x_i + 1) = \prod_{i=1}^{n} \left( \frac{2i}{n} \right) = \frac{2^n}{n^n} n! 
\]

**Consequence** Estimate for Newton-Cotes

\[
\left| \int_{-1}^{1} f(x)dx - Q_n(f) \right| \leq \frac{L}{(n+1)!} \frac{2^n}{n^n} n! = \frac{L}{(n+1)} \left( \frac{2}{n} \right)^n
\]

**Ratio** Chebyshev points to uniform grid

\[
\frac{n^n(n+1)}{(n+1)2^n \cdot 2^n} = \left( \frac{n}{4} \right)^n / n! \approx \left( \frac{e}{4} \right)^n \approx \left( \frac{2}{3} \right)^n
\]

Significant, but not overwhelming gain.

**Example** Trapezoidal Rule

\( n = 1, x_0 = -1, x_1 = +1 \)

\[
\int_{-1}^{1} f(x) = A_0 f(-1) + A_1 f(1) 
\]

\[
A_0 = \int_{-1}^{1} \ell_0(x) = \int_{-1}^{1} \frac{1}{2}(1 - x)dx = 1 - \frac{x^2}{4} \bigg|_{-1}^{+1} = 1 - \frac{1}{4} - \frac{1}{4} = \frac{1}{2} 
\]

\[
A_1 = \frac{1}{2} \quad \text{by symmetry} 
\]

\[
\int_{-1}^{1} |q_n(x)|dx = \int_{-1}^{1} |(x + 1)(x - 1)|dx = \int_{-1}^{1} (1 - x^2)dx = 2 - \frac{x^3}{3} \bigg|_{-1}^{+1} = 2 - \frac{1}{3} - \frac{1}{3} = \frac{4}{3} 
\]

\[
\left| \int_{-1}^{1} f(x)dx - \frac{1}{2}(f(-1) + f(+1)) \right| \leq \frac{L4}{23} = \frac{2}{3} 
\]

\[ \implies \text{after linear transformation} \]

\[
\left| \int_{a}^{b} f(x)dx - \frac{1}{2}[f(a) + f(b)] \right| \leq \frac{L(b-a)^3}{12} 
\]

**Composite trapezoidal rule**

\[
\left| \int_{a}^{b} f(x)dx - \left[ \frac{1}{2}[f(a) + f(b)] + \sum_{i=1}^{n-1} f(a + ih) \right] \right| \leq \frac{(b-a)^3}{12n^2} f^2(\xi) \leq \frac{(b-a)^3}{12n^2} L = \frac{(b-a)}{12} Lh^2 
\]
This is a second Order Method, i.e., doubling of \( n \) for even \( n \) developed by \( n \) leads to quartering of errors.

**Example** Composite Simpson

\[
\left| \int_a^b f(x)dx - \frac{h}{3} \left[ f(x_0) + 2 \sum_{i=2}^{n-2} f(x_{2i-2}) + 4 \sum_{i=1}^{n-2} f(x_{2i-1}) + f(x_n) \right] \right| \leq \frac{(b-a)}{180} L h^4
\]

**Remark** Composite Newton-Cotes methods have the advantage that, when \( n \) is doubled, all old values can be reused. Not possible for Chebyshev quadrature.

**Question:** Is it possible to design a quadrature using \( n + 1 \) points that is exact for polynomials of degree \( m \geq n \)?

**Answer:** Yes. By Gaussian quadrature yielding \( m = 2n + 1 \).

**Summary of I.1 and I.2**

\[
f \in C^{n,1}([a, b]), \quad q_n = \prod_{i=0}^{n} (x - x_i) \text{ monic}
\]

\[
\Rightarrow \int_a^b f(x)dx = \sum_{i=0}^{n} A_i f(x_i) + O \left( \frac{\|q_n\|_{\infty}}{(n+1)!} \right)
\]

\[
A_i = \int_a^b \ell_i(x)dx \quad \text{with} \quad \ell_i = \prod_{j \neq i}^{n} \frac{(x-x_j)}{(x_i-x_j)}
\]

**Exactness:** \( f \) polynomial of degree \( \leq n \Rightarrow L = 0 \)

\[
\Rightarrow \int_a^b f(x)dx = \sum_{i=0}^{n} A_i f(x_i).
\]

**Question:** Can the nodes = \( (n + 1) \) degrees of freedom be chosen such that exactness holds for all \( f \in \Pi_{2n+1} \), i.e., of degree \( 2n + 1 \)?

**Answer:** Yes. By Gaussian integration formula.

### 1.3 Gauss Integration

Generalized goal: For weight function \( 0 < w(x) \in C[a,b] \) (or piecewise continuous) try to achieve

\[
\int_a^b f(x)w(x)dx = \sum_{i=0}^{n} A_i f_i(x)
\]

for all polynomials \( f \) of degree \( \leq 2n + 1 \).

**First order case:** \( n = 0 \Rightarrow 2n + 1 = 1 \)

\[
\int_a^b w(x) = A_0 \cdot 1
\]

\[
\int_a^b x w(x)dx = A_0 x_0
\]

\[
x_0 = \frac{\int_a^b x w(x)dx}{\int_a^b w(x)dx} \in [a, b]
\]
Example:

\[
\begin{align*}
  w(x) &= e^x \Rightarrow \int_0^1 xe^x \, dx = x e^x \bigg|_0^1 - \int_0^1 e^x \, dx \\
  &= e - e + 1 = 1, \int_0^1 e^x \, dx = e - 1 \\
  x_0 &= e - 1 \approx 0.7
\end{align*}
\]

General Derivation:

\( \Pi[a, b] = \) polynomials on \([a, b]\) form inner product space w.r.t.

\[
\langle f, g \rangle = \int_a^b f(x)g(x)w(x)\, dx
\]

\( ||f||^2 = \int_a^b f(x)^2w(x)\, dx \)

\( \Pi_n[a, b] = \{ P \in \Pi[a, b] : \deg(P) \leq n \} \) is a Hilbert space of finite dimension \( \simeq \mathbb{R}^{n+1} \).

Monic representation:

\[
\begin{align*}
  f(x) &= \sum_{i=0}^n \varphi_i x^i, \quad g(x) = \sum_{i=0}^n \gamma_i x^i \\
  \langle f, g \rangle &= \bar{f}^\top H \bar{g}, \quad \bar{f} = (\varphi_i)_{i=0,...,n}, \quad \bar{g} = (\gamma_i)_{i=0,...,n} \\
  H &= (h_{ij}) = \left( \frac{1}{\int_0^1 x^i x^j \, dx} \right)_{j=0,...,n, i=0,...,n}
\end{align*}
\]

\( H \) is called Hilbert matrix and is terribly ill-conditioned.

Lemma 1.4 Gram-Schmidt orthogonalization

If \( \{ f_i \}_{0}^{n} \subset \Pi \) are linearly independent and nonzero, i.e. a basis of their span, then the recursion

\[
\begin{align*}
  v_0 &= f_0 \\
  v_1 &= f_1 - \langle v_0, f_1 \rangle v_0 \\
  v_k &= f_k - \sum_{j=1}^{k-1} \langle v_j, f_k \rangle v_j, \quad k = 2, ..., n
\end{align*}
\]

generates an orthogonal basis \( \{ v_i \}_{0}^{n} \), that is

\[
\langle v_j, v_k \rangle = \begin{cases} 
  ||v_j||^2 > 0 & \text{if } j = k, \\
  0 & \text{otherwise}.
\end{cases}
\]

Proof: Simple induction.

Proposition 1.5 For the generation of orthogonal polynomials with respect to the inner product \( \langle f, g \rangle = \int_a^b f(x)g(x)w(x)\, dx \), one obtains a 3-term recurrence in the following form:

Define \( P_{-1}(x) = 0, P_0(x) = 1 \), and

\[
\begin{align*}
  P_{k+1}(x) &= (x - \alpha_{k+1})P_k(x) - \gamma_k P_{k-1}(x) \quad k = 0, 1, ..., \\
  \text{where} \\
  \alpha_{k+1} &= \frac{\langle x P_k, P_k \rangle}{\langle P_k, P_k \rangle} \quad \text{for } k = 0, 1, ... \\
  \gamma_k &= \frac{\langle P_k, P_k \rangle}{\langle P_{k-1}, P_{k-1} \rangle} \quad \text{for } k = 1, 2, ...
\end{align*}
\]

Then the sequence \( P_0, P_1, P_2, ... \) generated in this form is orthogonal.
Proof: By induction in $k$

$$\langle P_1, P_0 \rangle = \langle x - \frac{\langle xP_0, P_0 \rangle}{\langle P_0, P_0 \rangle}, P_0, P_0 \rangle = \langle xP_0, P_0 \rangle - \langle xP_0, P_0 \rangle = 0$$

if orthogonality holds for $k - 1$, $k \geq 2$, then

$$\langle P_k, P_{k-1} \rangle = \langle xP_{k-1}, P_{k-1} \rangle - \langle \alpha_{k-1}P_{k-1}, P_{k-1} \rangle - \gamma_{k-1}\langle P_{k-1}, P_{k-2} \rangle$$

$$= \langle xP_{k-1}, P_{k-1} \rangle - \langle \alpha_{k-1}P_{k-1}, P_{k-1} \rangle = 0$$

$$\langle P_k, P_{k-2} \rangle = \langle xP_{k-1}, P_{k-2} \rangle - \gamma_{k-1}\langle P_{k-2}, P_{k-2} \rangle$$

$$= \langle P_{k-1}, xP_{k-2} \rangle - \langle P_{k-1}, P_{k-1} \rangle = 0$$

$$= \langle P_{k-1}, P_{k-1} + \alpha P_{k-2} + \gamma P_{k-3} \rangle - \langle P_{k-1}, P_{k-1} \rangle = 0$$

In addition, for $j < k - 2$ we have

$$\langle P_k, P_j \rangle = \langle (x - \alpha_k)P_{k-1} + \gamma_{k-1}P_{k-2}, P_j \rangle = \langle xP_{k-1}, P_j \rangle$$

$$= \langle P_{k-1}, xP_j \rangle = 0,$$

since $xP_j$ is in $\text{span}\{P_0, \ldots, P_{k-2}\}$, which is orthogonal to $P_{k-1}$ by inductive assumption.

Example: Legendre Polynomials:

$$[a, b] = [-1, 1], \ w(x) = 1,$$

$$\alpha_1 = \int_{-1}^{1} xdx / \int_{-1}^{1} 1dx = 0, P_1 = x$$

$$\alpha_2 = \int_{-1}^{1} x^3dx / \int_{-1}^{1} x^2 dx = 0$$

$$\gamma_1 = \int_{-1}^{1} x^2 \cdot 1 / \int_{-1}^{1} 1dx = \frac{1}{3} \Rightarrow P_2 = x^2 - \frac{1}{3}$$

and so on

$$P_3 = x^3 - \frac{3}{5} x, \ P_4 = x^4 - \frac{6}{7} x^2 + \frac{3}{35}$$

Theorem 1.6 Let $0 < w(x) \in C[a, b]$, and consider the quadrature $Q_n(f) = \sum_{i=0}^{n} A_i f(x_i)$, where $x_i, 0 \leq i \leq n$, are defined as the $n + 1$ zeros of the orthogonal polynomial $P_{n+1}$ (orthogonal to $\Pi_n[a, b]$). Then $Q_n$ is exact for all polynomials in $\Pi_{2n+1}[a, b]$. The coefficients $A_i$ are given by

$$A_i = \int_{a}^{b} w(x)\ell_i^2(x)dx > 0 \text{ for } i = 0, 1, \ldots, n$$

and sum to $\int_{a}^{b} w(x)dx$.

Proof: By polynomial division

$$f \in \Pi_{2n+1}[a, b] = P_n(x)S(x) + R(x)$$
with \( \deg(S(x)) \leq n \) and \( \deg(R(x)) \leq n \)

\[
\int_a^b f(x)w(x)dx = \int_a^b q_n(x)w(x)S(x)dx + \int_a^b R(x)w(x)dx = Q_n(R) = Q_n(f).
\]

The squared Lagrangians \( l_i^2(x) \) have degree \( 2n \) and are therefore integrable exactly. The exactness on \( f(x) = 1 \) implies that \( \sum_{i=0}^n A_i = \int_a^b w(x)dx \).

**Remark 1.7** The theorem holds also for discontinuous \( w \), as long as the inner product is well defined for polynomials.

**Corollary 1.8** When \( w \equiv 1 \) for \( f \in C([a,b]) \) and \( P \in \Pi_{2n+1}[a,b] \), then the Gauss quadrature \( Q_n \) yields

\[
\left| \int_a^b f(x)dx - Q_n(f) \right| \leq 2 \cdot (b-a)\|f - P\|_\infty
\]

**Proof:**

\[
\left| \int_a^b (f(x)-P(x))dx - Q_n(f-P) \right| \leq (b-a)\|f - P\|_\infty + \sum_{i=0}^n A_i |f(x_i)-P(x_i)| \leq 2(b-a)\|f - P\|_\infty.
\]

**Question:** Which are candidates for \( P \)?

Take Chebychev of order \((2n+1)\) when \( f \in C^{2n+1,1} \), then we have

\[
\|f - P\|_\infty \leq \frac{L}{(2n+2)!2^{2n+1}},
\]

by Lemma I.2 for \([a,b] = [-1,1] \).

Resulting estimate for quadrature

\[
\left| \int_a^b f(x)dx - Q_n(f) \right| \leq \frac{L}{(2n+2)!2^{2n-1}}
\]

corresponds to that for ”simple” quadrature at \((2n+1)\) Chebyshev points.

**Sharp estimation by Hermite interpolation**

**Proposition 1.9** If \( f \in C^{2n+2}([a,b]) \) and \( a_0 \leq x_0 < \cdots < x_{n-1} < x_n \leq b \), then there exists a unique polynomial of degree \( 2n+1 \) such that

\[
\begin{align*}
P(x_i) &= f(x_i) \\
P'(x_i) &= f'(x_i)
\end{align*}
\]

for \( 0 \leq i \leq n \)

and for all \( x \in [a,b] \)

\[
f(x) - P(x) = \frac{f^{(2n+2)}(\xi)}{(2n+2)!} \prod_{i=0}^n (x-x_i)^2
\]

for some mean-value \( a < \xi < b \).
Corollary 1.10 The error of Gaussian quadrature is given by
\[
\int_a^b f(x)w(x)dx - Q_n(f) = \frac{f^{(2n+2)}(\xi')}{(2n+2)!}\int_a^b \prod_{i=0}^n (x-x_i)^2w(x)dx.
\]
If \(w(x)=1\) and \([a,b]=[-1,1]\), we have by Lemma 1.3
\[
\int_a^b \prod_{i=0}^n (x-x_i)^2dx \geq \frac{1}{2^{2n}}
\]
i.e., we gain at most a factor of 2 compared to simple argument.

Calculating Gauss Points and Weights
One possibility is to compute the roots of the polynomial \(P_{n+1}(x)\) by Newton method. Slick alternative: Reformulation as symmetric eigenvalue problem.

Recursion \(P_{-1}(x) = 0, P_0(x) = 1, \) and
\[
P_{k+1}(x) = (x-\alpha_{k+1})P_k(x) - \gamma_k P_{k-1}(x) \quad k = 0, 1, \ldots,
\]
where
\[
\alpha_{k+1} = \frac{\langle xP_k, P_k \rangle}{\langle P_k, P_k \rangle} \text{ for } k = 0, 1, \ldots
\]
\[
\gamma_k = \frac{\langle P_k, P_{k-1} \rangle}{\langle P_{k-1}, P_{k-1} \rangle} \text{ for } k = 1, 2, \ldots,
\]
can be rewritten using the normalized polynomials \(\hat{P}_k = \frac{P_k}{\|P_k\|}\) (i.e. replacing \(P_k = \hat{P}_k \|P_k\|\)) and dividing the recurrence by \(\|P_k\|\) as
\[
\hat{P}_{-1}(x) = 0, \hat{P}_0(x) = 1/(\beta_0) \text{ with } \beta_0 = \int_a^b w(x)dx, \text{ and}
\]
\[
\sqrt{\beta_{k+1}} \hat{P}_{k+1}(x) = (x-\alpha_{k+1})\hat{P}_k(x) - \sqrt{\beta_k} \hat{P}_{k-1}(x) \quad k = 0, 1, \ldots,
\]
where
\[
\alpha_{k+1} = \langle x\hat{P}_k, \hat{P}_k \rangle = \frac{\langle xP_k, P_k \rangle}{\langle P_k, P_k \rangle} \text{ for } k = 0, 1, \ldots
\]
\[
\beta_k = \frac{\langle P_k, P_{k-1} \rangle}{\langle P_{k-1}, P_{k-1} \rangle} \text{ for } k = 1, 2, \ldots.
\]
Thus we have that \(\alpha_{k+1}\) is the same as in the orthogonal 3-term recurrence, and now \(\beta_k = \gamma_k\). The 3-term recurrence can be written in matrix form as following

\[
\begin{bmatrix}
P_0(x) \\
P_1(x)
\end{bmatrix}
= \begin{bmatrix}
\alpha_1 & \sqrt{\beta_1} \\
\sqrt{\beta_1} & \alpha_2 & \sqrt{\beta_2} \\
& \sqrt{\beta_2} & \alpha_3 & \sqrt{\beta_3} \\
& & & \sqrt{\beta_k} & \alpha_{k+1}
\end{bmatrix}
\begin{bmatrix}
P_0(x) \\
P_1(x)
\end{bmatrix}
+ \begin{bmatrix}
0 \\
\sqrt{\beta_{k+1}} \hat{P}_{k+1}(x)
\end{bmatrix}
\]

\[
\begin{bmatrix}
P_0(x) \\
P_1(x)
\end{bmatrix}
= \begin{bmatrix}
\alpha_1 & \sqrt{\beta_1} \\
\sqrt{\beta_1} & \alpha_2 & \sqrt{\beta_2} \\
& \sqrt{\beta_2} & \alpha_3 & \sqrt{\beta_3} \\
& & & \sqrt{\beta_k} & \alpha_{k+1}
\end{bmatrix}
\begin{bmatrix}
P_0(x) \\
P_1(x)
\end{bmatrix}
\]

\[
T v^{(i)} = v^{(i)} x_i \text{ with } v^{(i)} = \begin{pmatrix} P_0(x_i) \\ \vdots \\ P_k(x_i) \end{pmatrix} \in \mathbb{R}^{n+1}.
\]
All eigenvalues are real since \(T\) is a symmetric matrix.
**Proposition 1.11** (Golub and Welsh 1968)
The $(n + 1)$ roots $x_i$ for $i = 0, \ldots, n$ of $P_{n+1}$ are distinct, real and can be calculated by eigenvalue decomposition of $T$. Weights $A_i$ are given by

$$A_i = \left( v_1^{(i)} \right)^2 \int_a^b w(x) dx$$

where $v_1^{(i)}$ is the first component of the normalized eigenvector.

**Proof:** Golub and Welsh

**Remark:** Eigenvalue decomposition is computed by OR algorithm on $T$. Symmetry and boundedness is maintained so that everything can happen in $0(n)$ operations per iteration.