

Math 270C: Numerical Mathematics (Part C)
LECTURE NOTES

Bo Li
Department of Mathematics
University of California, San Diego

June 11, 2007

Warning!

- While being expanded with the addition of new material and being carefully polished continuously, these notes may still contain many typos and mistakes.
- These notes are mainly for the graduate course Math 270C, Spring quarter, 2007, at UC San Diego. Any individuals are however welcome to use these notes for personal studies and classes. As always, any comments are very much appreciated.
- Please do not copy these notes partially or entirely for publication without permission of the author. Please do not distribute these notes.

Acknowledgments

Many thanks to Yi Ling Ng for typing out the first version of the latex file of these notes, and to the graduate students of Math 270C, Spring quarter, 2007, for their valuable comments on these notes.

Contents

1	Polynomial Approximation	5
1.1	The Weierstrass Theorem	5
1.2	Best Uniform Approximation	8
1.3	Chebyshev Polynomials	14
1.4	Uniform Approximation by Trigonometric Polynomials	17
1.5	Modulus of Continuity and Jackson's Theorems	17
1.6	Least-Squares Approximation	18
1.7	Orthogonal Polynomials	21
1.8	More Properties of Orthogonal Polynomials	25
1.9	Legendre Polynomials	30
	Exercises	33
2	Polynomial Interpolation	41
2.1	Lagrange Interpolation	41
2.2	Newton's Formula and Divided Differences	46
2.3	Peano Kernel and the Remainder Theorem	52
2.4	Hermite Interpolation and Divided Differences with Repeated Points	56
2.5	Convergence of Interpolation Polynomials	58
2.6	Splines	62
2.7	Trigonometric Interpolation and Fast Fourier Transforms	62
	Exercises	62
3	Numerical Integration	69
3.1	The Basics	69
3.2	Interpolatory Quadrature	74
3.3	Peano Kernel and Error Representation	81
3.4	Euler-Maclaurin Formula and Romberg Algorithm	82
3.5	Weighted Gaussian Quadrature	82
3.6	Convergence of Sequences of Numerical Quadrature	88
3.7	Adaptive Method	90
	Exercises	90

Chapter 1

Polynomial Approximation

1.1 The Weierstrass Theorem

Let $a, b \in \mathbb{R}$ with $a < b$. Let $C[a, b]$ denote the set of all continuous, real-valued functions on the closed interval $[a, b]$. Let \mathcal{P} denote the set of all real polynomials. For each integer $n \geq 0$, let \mathcal{P}_n denote the set of all real polynomials of degree less than or equal to n .

Theorem 1.1 (The First Weierstrass Approximation Theorem). *Let $f \in C[a, b]$. For any $\varepsilon > 0$, there exists $p \in \mathcal{P}$ such that*

$$|f(x) - p(x)| < \varepsilon \quad \forall x \in [a, b]. \quad (1.1)$$

To prove Theorem 1.1, we define for any given $f \in C[0, 1]$

$$(B_0 f)(x) = f(0) \quad \forall x \in [0, 1],$$
$$(B_n f)(x) = \sum_{k=0}^n f\left(\frac{k}{n}\right) \binom{n}{k} x^k (1-x)^{n-k} \quad \forall x \in [0, 1] \quad \forall n \geq 1,$$

where

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

For each $n \geq 0$, we call $B_n f$ the n th Bernstein polynomial of $f \in C[0, 1]$. Clearly, $B_n f \in \mathcal{P}_n$ for each n .

Lemma 1.2 (Properties of Bernstein polynomials).

(1) For each $n \geq 0$, $B_n : C[0, 1] \rightarrow \mathcal{P}_n$ is linear, i.e.,

$$B_n(f + g) = B_n f + B_n g \quad \forall f, g \in C[0, 1],$$
$$B_n(\alpha f) = \alpha B_n f \quad \forall \alpha \in \mathbb{R} \quad \forall f \in C[0, 1].$$

(2) For each $n \geq 0$, $B_n : C[0, 1] \rightarrow \mathcal{P}_n$ is non-negative, i.e., $(B_n f)(x) \geq 0$ for any $x \in [0, 1]$ provided that $f(x) \geq 0$ for any $x \in [0, 1]$.

(3) Let $p_i(x) = x^i$ ($i = 1, 2, 3$). Then,

$$(B_n p_0)(x) = p_0(x) \quad \forall x \in [0, 1] \quad \forall n \geq 0, \quad (1.2)$$

$$(B_n p_1)(x) = p_1(x) \quad \forall x \in [0, 1] \quad \forall n \geq 1, \quad (1.3)$$

$$(B_n p_2)(x) = p_2(x) + \frac{1}{n}x(1-x) \quad \forall x \in [0, 1] \quad \forall n \geq 2. \quad (1.4)$$

Proof. Part (1) and Part (2) are obvious. To prove Part (3), we use the binomial formula:

$$(\alpha + \beta)^n = \sum_{k=0}^n \binom{n}{k} \alpha^k \beta^{n-k} \quad \forall \alpha, \beta \in \mathbb{R}. \quad (1.5)$$

Let $x \in [0, 1]$. We have by (1.5) that for $n \geq 1$

$$(B_n p_0)(x) = \sum_{k=0}^n \binom{n}{k} x^k (1-x)^{n-k} = (x + (1-x))^n = 1 = p_0(x).$$

By definition, $(B_0 p_0)(x) = p_0(x) = 1$ for all $x \in [0, 1]$. Thus, (1.2) is true.

Let $n \geq 1$. Using (1.5), we obtain for any $x \in [0, 1]$ that

$$\begin{aligned} (B_n p_1)(x) &= \sum_{k=0}^n p_1 \left(\frac{k}{n} \right) \binom{n}{k} x^k (1-x)^{n-k} \\ &= \sum_{k=0}^n \frac{k}{n} \frac{n!}{k!(n-k)!} x^k (1-x)^{n-k} \\ &= x \sum_{k=1}^n \frac{(n-1)!}{(k-1)!((n-1)-(k-1))!} x^{k-1} (1-x)^{(n-1)-(k-1)} \\ &\stackrel{j=k-1}{=} x \sum_{j=0}^{n-1} \binom{n-1}{j} x^{n-1-j} (1-x)^{n-1-j} \\ &= x(x + (1-x))^{n-1} \\ &= p_1(x). \end{aligned}$$

This is (1.3). The proof of (1.4) is left as an exercise. □

Proof of the First Weierstrass Approximation Theorem. We consider first the special case $[a, b] = [0, 1]$ and then the general case.

Case 1. $[a, b] = [0, 1]$.

Step 1. Let $M = \max_{0 \leq x \leq 1} |f(x)|$. Then,

$$-2M \leq f(x) - f(y) \leq 2M \quad \forall x, y \in [0, 1]. \quad (1.6)$$

Since $f \in C[0, 1]$, f is uniformly continuous on $[0, 1]$. Thus, there exists $\delta > 0$ such that

$$-\frac{\varepsilon}{2} < f(x) - f(y) < \frac{\varepsilon}{2} \quad \forall x, y \in [0, 1] \text{ with } |x - y| < \delta. \quad (1.7)$$

We claim that

$$-\frac{\varepsilon}{2} - \frac{2M}{\delta^2}(x - y)^2 \leq f(x) - f(y) \leq \frac{\varepsilon}{2} + \frac{2M}{\delta^2}(x - y)^2 \quad \forall x, y \in [0, 1]. \quad (1.8)$$

In fact, if $|x - y| < \delta$, then (1.7) implies (1.8). If $|x - y| \geq \delta$, then $((x - y)/\delta)^2 \geq 1$, and hence, (1.6) implies (1.8).

Step 2. Fix y in (1.8). Apply B_n ($n \geq 1$) to each side of (1.8) as a continuous function of x . Using the properties of B_n (cf. Lemma 1.2), we obtain

$$-\frac{\varepsilon}{2} - \frac{2M}{\delta^2}(B_n(x - y)^2)(x) \leq (B_n f)(x) - f(y) \leq \frac{\varepsilon}{2} + \frac{2M}{\delta^2}(B_n(x - y)^2)(x). \quad (1.9)$$

Since

$$(x - y)^2 = x^2 - 2xy + y^2 = p_2(x) - 2yp_1(x) + y^2p_0(x),$$

we have again by Lemma 1.2 that

$$\begin{aligned} (B_n(x - y)^2)(x) &= (B_n p_2)(x) - 2y(B_n p_1)(x) + y^2(B_n p_0)(x) \\ &= x^2 + \frac{1}{n}x(1 - x) - 2yx + y^2 \\ &= (x - y)^2 + \frac{1}{n}x(1 - x). \end{aligned}$$

This and (1.9) lead to

$$-\frac{\varepsilon}{2} - \frac{2M}{\delta^2} \left[(x - y)^2 + \frac{x(1 - x)}{n} \right] \leq (B_n f)(x) - f(y) \leq \frac{\varepsilon}{2} + \frac{2M}{\delta^2} \left[(x - y)^2 + \frac{x(1 - x)}{n} \right] \quad \forall x, y \in [0, 1]. \quad (1.10)$$

Step 3. Setting $y = x$ in (1.10), we obtain

$$|(B_n f)(x) - f(x)| \leq \frac{\varepsilon}{2} + \frac{2M}{\delta^2} \frac{x(1 - x)}{n} \leq \frac{\varepsilon}{2} + \frac{2M}{\delta^2 n} \quad \forall x \in [0, 1].$$

Choosing $n > 4M/(\delta^2\varepsilon)$ and $p = B_n f \in \mathcal{P}$, we then obtain (1.1) for $[a, b] = [0, 1]$.

Case 2. For general $[a, b]$, define $g(t) = f((b - a)t + a)$ ($0 \leq t \leq 1$). Then, $g \in C[0, 1]$. Hence, by what has been proved in Case 1, there exists $q \in \mathcal{P}$ such that

$$|g(t) - q(t)| < \varepsilon \quad \forall t \in [0, 1].$$

Let $p(x) = q((x - a)/(b - a))$. Then, $p \in \mathcal{P}$, and

$$|f(x) - p(x)| = |g(t) - q(t)| < \varepsilon \quad \forall x \in [a, b].$$

This is (1.1). □

We define for any $f \in C[a, b]$,

$$\|f\|_{C[a,b]} = \max_{a \leq x \leq b} |f(x)|.$$

When no confusion arises, we also write $\|f\|$ instead of $\|f\|_{C[a,b]}$. One easily verifies that $\|\cdot\|$ is a norm of the vector space $C[a, b]$, i.e., the following hold:

- (1) $\|f\| \geq 0 \quad \forall f \in C[a, b]$.
 $\|f\| = 0$ if and only if $f(x) = 0 \quad \forall x \in [a, b]$;
- (2) $\|\alpha f\| = |\alpha| \|f\| \quad \forall \alpha \in \mathbb{R} \quad \forall f \in C[a, b]$;
- (3) $\|f + g\| \leq \|f\| + \|g\| \quad \forall f, g \in C[a, b]$.

The last inequality is called the *triangle inequality*. It implies that

$$|\|f\| - \|g\|| \leq \|f - g\| \quad \forall f, g \in C[a, b]. \quad (1.11)$$

This norm is called the *maximum norm* or $C[a, b]$ -*norm*.

With the notion of maximum norm, the Weierstrass Theorem states exactly that, for any $f \in C[a, b]$ and any $\varepsilon > 0$, there exists $p \in \mathcal{P}$ such that $\|f - p\| < \varepsilon$. Equivalently, for any $f \in C[a, b]$, there exists a sequence of polynomials $\{p^{(k)}\}_{k=1}^{\infty}$ such that $p^{(k)} \rightarrow f$ in $C[a, b]$, i.e., $\|f - p^{(k)}\| \rightarrow 0$ as $k \rightarrow \infty$.

The above proof of the First Weierstrass Approximation Theorem using the Bernstein polynomials can be extended to the proof of the following Bohman-Korovkin Theorem for more general sequences of linear, non-negative operators; and clearly the Bohman-Korovkin Theorem and the properties of Bernstein polynomials imply the First Weierstrass Approximation Theorem:

Theorem 1.3 (Bohman-Korovkin). *For each integer $n \geq 1$, let $L_n : C[a, b] \rightarrow \mathcal{P}_n$ be an operator. Assume:*

- (1) *Each $L_n : C[a, b] \rightarrow \mathcal{P}_n$ is linear;*
- (2) *Each $L_n : C[a, b] \rightarrow \mathcal{P}_n$ is non-negative;*
- (3) *$L_n p_k \rightarrow p_k$ in $C[a, b]$ for $k = 0, 1, 2$, where $p_k(x) = x^k$.*

Then $\|L_n f - f\| \rightarrow 0$ for any $f \in C[a, b]$. □

1.2 Best Uniform Approximation

We denote for any $f \in C[a, b]$

$$E_n(f) := \inf_{p \in \mathcal{P}_n} \|f - p\| \quad \forall n \geq 0. \quad (1.12)$$

Here, the norm $\|\cdot\|$ is the $C[a, b]$ -norm. Clearly, $0 \leq E_n(f) \leq \|f\|$ for all $n \geq 0$. Moreover, since $\mathcal{P}_0 \subset \cdots \subset \mathcal{P}_n \subset \cdots$, we have $E_0(f) \geq \cdots \geq E_n(f) \geq \cdots$. The First Weierstrass Approximation Theorem then implies the following:

Proposition 1.4. *We have*

$$\lim_{n \rightarrow \infty} E_n(f) = 0 \quad \forall f \in C[a, b]. \quad \square$$

Definition 1.5 (Best uniform approximation). *A best uniform approximation of a given function $f \in C[a, b]$ in \mathcal{P}_n is a polynomial $p_n \in \mathcal{P}_n$ that satisfies*

$$\|f - p_n\| = \min_{q_n \in \mathcal{P}_n} \|f - q_n\|. \quad (1.13)$$

Since (1.13) can be written as

$$\max_{a \leq x \leq b} |f(x) - p_n(x)| = \min_{q_n \in \mathcal{P}_n} \max_{a \leq x \leq b} |f(x) - q_n(x)|,$$

a best uniform approximation is also called a *minimax approximation*.

Theorem 1.6 (Existence of best uniform approximations). *For any $f \in C[a, b]$ and any integer $n \geq 0$, there exists a best uniform approximation of f in \mathcal{P}_n .*

Proof. Let $f \in C[a, b]$ and fix an integer $n \geq 0$. For any $c = (c_0, \dots, c_n) \in \mathbb{R}^{n+1}$, we associate with a unique polynomial $p_c \in \mathcal{P}_n$ by $p_c(x) = \sum_{k=0}^n c_k x^k$. Define $F : \mathbb{R}^{n+1} \rightarrow \mathbb{R}$ by

$$F(c) = \|f - p_c\| = \max_{a \leq x \leq b} \left| f(x) - \sum_{k=0}^n c_k x^k \right|.$$

Clearly, the assertion of the theorem is equivalent to the existence of $c \in \mathbb{R}^{n+1}$ such that

$$F(c) = \min_{d \in \mathbb{R}^{n+1}} F(d). \quad (1.14)$$

Step 1. We show that the function $F : \mathbb{R}^{n+1} \rightarrow \mathbb{R}$ satisfies the following properties:

- (1) $F : \mathbb{R}^{n+1} \rightarrow \mathbb{R}$ is continuous;
- (2) $F(c) \geq 0$ for any $c \in \mathbb{R}^{n+1}$;
- (3) $F(c) \rightarrow +\infty$ as $c \rightarrow \infty$.

By (1.11), we have

$$\begin{aligned} |F(c) - F(d)| &= \left| \|f - p_c\| - \|f - p_d\| \right| \leq \|p_d - p_c\| \\ &= \max_{a \leq x \leq b} \left| \sum_{k=0}^n (d_k - c_k) x^k \right| \leq \sum_{k=0}^n |d_k - c_k| \max\{|a|^k, |b|^k\} \\ &\rightarrow 0 \quad \text{as } d \rightarrow c \text{ in } \mathbb{R}^{n+1}. \end{aligned}$$

This proves Part (1). Part (2) is clearly true.

To prove Part (3), we define $G : \mathbb{R}^{n+1} \rightarrow \mathbb{R}$ by $G(c) = \|p_c\|$ for any $c \in \mathbb{R}^{n+1}$. As proved in Part (1), $G : \mathbb{R}^{n+1} \rightarrow \mathbb{R}$ is continuous. Let

$$\mathbb{S}^n = \{c \in \mathbb{R}^{n+1} : \|c\| = 1\},$$

where

$$\|c\| = \sqrt{\sum_{k=0}^n c_k^2} \quad \forall c = (c_0, \dots, c_n) \in \mathbb{R}^{n+1}.$$

Clearly, \mathbb{S}^n is a closed and bounded subset, and hence a compact subset, of \mathbb{R}^{n+1} . Moreover, $G(c) > 0$ for any $c \in \mathbb{S}^n$. The continuity of $G : \mathbb{R}^{n+1} \rightarrow \mathbb{R}$ then implies that

$$\mu := \inf_{c \in \mathbb{S}^n} G(c) > 0. \quad (1.15)$$

Now, let $c \in \mathbb{R}^{n+1}$ with $c \neq 0$. Set $\hat{c} := (1/\|c\|)c \in \mathbb{S}^n$. We have

$$p_c(x) = \sum_{k=0}^n c_k x^k = \|c\| \sum_{k=0}^n \frac{c_k}{\|c\|} x^k = \|c\| p_{\hat{c}}(x) \quad \forall x \in [a, b],$$

i.e., $p_c = \|c\| p_{\hat{c}}$. Therefore, by (1.11) and (1.15), we have

$$\begin{aligned} F(c) &\geq \|p_c\| - \|f\| = \| \|c\| p_{\hat{c}} \| - \|f\| \\ &= \|c\| G(p_{\hat{c}}) - \|f\| \geq \|c\| \mu - \|f\| \rightarrow +\infty \quad \text{as } \|c\| \rightarrow \infty, \end{aligned}$$

proving Part (3).

Step 2. Let $m = \inf_{c \in \mathbb{R}^{n+1}} F(c) \geq 0$. Since $\lim_{c \rightarrow \infty} F(c) = +\infty$, there exists $R > 0$ such that $F(c) \geq m + 1$ if $\|c\| > R$. Therefore, $m = \inf_{\|c\| \leq R} F(c)$. Since F is continuous on the bounded and closed subset $B_R := \{c \in \mathbb{R}^{n+1} : \|c\| \leq R\}$, it attains its minimum over B_R , and hence over the entire \mathbb{R}^{n+1} . \square

Theorem 1.7 (The Chebyshev Alternation Theorem). *Let $f \in C[a, b]$ and $f \notin \mathcal{P}_n$. Then, $p \in \mathcal{P}_n$ is a best uniform approximation if and only if $f - p$ achieves its maximum magnitude at $n + 2$ points with alternating signs, i.e., there exist $n + 2$ points x_k ($1 \leq k \leq n + 2$) with $a \leq x_1 < \dots < x_{n+2} \leq b$ such that*

$$\begin{aligned} |f(x_k) - p(x_k)| &= \|f - p\|, \quad 1 \leq k \leq n + 2, \\ [f(x_k) - p(x_k)][f(x_{k+1}) - p(x_{k+1})] &< 0, \quad k = 1, \dots, n + 1. \end{aligned}$$

Example. Consider $f(x) = x^2$ with $x \in [0, 1]$ and $n = 1$. Let $p(x) = x - 1/8$. So, $p \in \mathcal{P}_1$. Let

$$g(x) = f(x) - p(x) = x^2 - x + \frac{1}{8}.$$

It is easy to show that $|g(x)|$ attains its maximum value $\|g\| = 1/8$ at $x = 0, 1/2, 1$. Also,

$$g(0) = -g\left(\frac{1}{2}\right) = g(1) = \frac{1}{8} = \|g\|.$$

Thus, by the Chebyshev Alternation Theorem, $p_1 \in \mathcal{P}_1$ is a best uniform approximation of $f(x) = x^2$ on $[0, 1]$, i.e.,

$$\max_{0 \leq x \leq 1} |f(x) - p(x)| = \max_{0 \leq x \leq 1} \left| x^2 - x + \frac{1}{8} \right| \leq \max_{0 \leq x \leq 1} |x^2 - q(x)| \quad \forall q \in \mathcal{P}_1.$$

Equivalently,

$$\max_{0 \leq x \leq 1} \left| x^2 - x + \frac{1}{8} \right| = \min_{\alpha, \beta \in \mathbb{R}} \max_{0 \leq x \leq 1} |x^2 + \alpha x + \beta|.$$

Definition 1.8 (Change of sign). A function $g : (a, b) \rightarrow \mathbb{R}$ changes its sign at a point $z \in (a, b)$, if there exists $\varepsilon > 0$ with $(z - \varepsilon, z + \varepsilon) \subset (a, b)$ such that either the following hold true:

- (1) $g(x) \geq 0$ for all $x \in (z - \varepsilon, z)$ and $g(x) \leq 0$ for all $x \in (z, z + \varepsilon)$;
- (2) both one-sided limits $g(z-)$ and $g(z+)$ exist and $g(z-) \neq g(z+)$,

or the following hold true:

- (1) $g(x) \leq 0$ for all $x \in (z - \varepsilon, z)$ and $g(x) \geq 0$ for all $x \in (z, z + \varepsilon)$;
- (2) both one-sided limits $g(z-)$ and $g(z+)$ exist and $g(z-) \neq g(z+)$.

Proof of the Chebyshev Alternation Theorem. It is clear that the Chebyshev Alternation Theorem is equivalent to the following statement:

Let $f \in C[a, b]$ but $f \notin \mathcal{P}_n$. Then, the zero polynomial $0 \in \mathcal{P}_n$ is a best uniform approximation of f in \mathcal{P}_n if and only if f achieves its maximum magnitude at $n + 2$ points in $[a, b]$ with alternating signs.

Therefore, we need only to prove this statement. We divide our proof into two parts.

Part 1. The “if” part. If 0 were not the best uniform approximation in \mathcal{P}_n of f , then there would exist $p \in \mathcal{P}_n$ such that

$$0 < \|f - p\| < \|f - 0\| = \|f\|. \tag{1.16}$$

Let $a \leq x_1 < \dots < x_{n+2} \leq b$ be such that

$$\begin{aligned} |f(x_k)| &= \|f\|, & k &= 1, \dots, n+2, \\ f(x_k)f(x_{k+1}) &< 0 & k &= 1, \dots, n+1. \end{aligned}$$

It follows from (1.16) that

$$|p(x_k) - f(x_k)| \leq \|p - f\| < \|f\| = |f(x_k)|.$$

Consequently,

$$\text{sign } p(x_k) = \text{sign } (p(x_k) - f(x_k) + f(x_k)) = \text{sign } f(x_k), \quad k = 1, \dots, n+2,$$

where $\text{sign } z = z/|z|$ for any nonzero $z \in \mathbb{R}$ and $\text{sign } 0 = 0$. Thus, p changes its sign at $n + 2$ points. Hence, p has at least $n + 1$ roots. But $p \in \mathcal{P}_n$. So, $p = 0$. By (1.16), $\|0 - f\| < \|f\|$. This is a contradiction.

Part 2. The “only if” part. Assume the assertion were not true. Then, there would exist $m + 1$ points with $0 \leq m \leq n$ such that f achieves its maximum magnitude at these points with alternating signs. If $m = 0$, then f would have the same sign at all its points of maximum magnitude. In this case, there would exist a nonzero constant c such that $\|f - c\| < \|f\|$. This contradicts the fact that 0 is the best uniform approximation of f in \mathcal{P}_n . Thus, $1 \leq m \leq n$. We shall construct a polynomial in \mathcal{P}_n that would be closer to f than 0 with respect to the norm $\|\cdot\|$.

Step 1. Let $Z = \{x \in [a, b] : |f(x)| = \|f\|\}$. Without loss of generality, we may assume that $f > 0$ at the smallest number of Z . Such smallest number exists, since Z is closed in $[a, b]$, a consequence of the continuity of f . Since $\text{sign } f$ alternates at $m + 1$ points in Z , there exist ξ_k ($1 \leq k \leq m$) with $a \leq \xi_1 < \cdots < \xi_m \leq b$ such that $\text{sign } f$ alternates in $Z \cap [\xi_k, \xi_{k+1}]$ ($k = 1, \dots, m$). Define

$$p(x) = (\xi_1 - x) \cdots (\xi_m - x).$$

Clearly, $p \in \mathcal{P}_n$, and

$$\text{sign } p(x) = \text{sign } f(x) \quad \forall x \in Z. \quad (1.17)$$

Let $x \in Z$. Since $|f(x)| = \|f\| > 0$, (1.17) implies that $p(x) \neq 0$. Since Z is closed in $[a, b]$,

$$\min_{x \in Z} |p(x)| > 0. \quad (1.18)$$

In particular, this implies that $\|p\| > 0$.

Step 2. We show that

$$\left\| f - \frac{\varepsilon}{\|p\|} p \right\| < \|f\| \quad (1.19)$$

for small $\varepsilon > 0$.

Let

$$\delta = \min_{x \in Z} \frac{p(x)f(x)}{\|p\|} = \min_{x \in Z} \frac{|p(x)||f(x)|}{\|p\|} = \frac{\|f\|}{\|p\|} \min_{x \in Z} |p(x)|.$$

By (1.18), $\delta > 0$. Define

$$A = \left\{ x \in [a, b] : \frac{p(x)f(x)}{\|p\|} > \frac{\delta}{2} \right\} \quad \text{and} \quad B = \left\{ x \in [a, b] : \frac{p(x)f(x)}{\|p\|} \leq \frac{\delta}{2} \right\}.$$

Clearly, $[a, b] = A \cup B$, A and B are disjoint, and B is closed in $[a, b]$. If $x \in A$, then

$$\left| f(x) - \frac{\varepsilon}{\|p\|} p(x) \right|^2 = |f(x)|^2 - 2\varepsilon \frac{p(x)f(x)}{\|p\|} + \varepsilon^2 \left(\frac{p(x)}{\|p\|} \right)^2$$

$$\leq \|f\|^2 - \varepsilon\delta + \varepsilon^2 = \|f\|^2 - \frac{\delta^2}{4} \quad \text{for } 0 < \varepsilon < \frac{\delta}{2}. \quad (1.20)$$

If $x \in B$, then by the definition of δ and B ,

$$\frac{p(x)f(x)}{\|p\|} \leq \frac{\delta}{2} < \delta \leq \frac{p(y)f(y)}{\|p\|} \quad \forall y \in Z.$$

This implies that $x \notin Z$, and hence $|f(x)| < \|f\|$. Consequently, since f is continuous and B is closed in $[a, b]$, we obtain that

$$M := \max_{x \in B} |f(x)| < \|f\|.$$

Therefore, for $0 < \varepsilon < \|f\| - M$,

$$\left| f(x) - \frac{\varepsilon}{\|p\|} p(x) \right| \leq |f(x)| + \left| \frac{\varepsilon}{\|p\|} p(x) \right| \leq M + \varepsilon < \|f\| \quad \forall x \in B. \quad (1.21)$$

It therefore follows from (1.20) and (1.21) that

$$\left\| f - \frac{\varepsilon}{\|p\|} p \right\| \leq \|f\|, \quad \text{if } 0 < \varepsilon < \min \left(\frac{\delta}{2}, \|f\| - M \right).$$

This contradicts the fact that 0 is the best uniform approximation in \mathcal{P}_n of f . \square

Theorem 1.9 (Uniqueness of best uniform approximations). *For any $f \in C[a, b]$ and integer $n \geq 0$, the best approximation of f in \mathcal{P}_n is unique.*

Proof. Without loss of generality, we may assume $f \notin \mathcal{P}_n$. Assume that $p, q \in \mathcal{P}_n$ are best uniform approximations of f in \mathcal{P}_n , i.e.,

$$\|f - p\| = \|f - q\| = E_n(f),$$

where $E_n(f)$ is defined in (1.12). Let $r = (p + q)/2 \in \mathcal{P}_n$. Then,

$$E_n(f) \leq \|f - r\| = \left\| \frac{1}{2}(f - p) + \frac{1}{2}(f - q) \right\| \leq \frac{1}{2}\|f - p\| + \frac{1}{2}\|f - q\| = E_n(f).$$

Hence, $r \in \mathcal{P}_n$ is also a best uniform approximation of f in \mathcal{P}_n .

By the Chebyshev Alternation Theorem, there exist $n + 2$ points x_k ($k = 1, \dots, n + 2$) with $a \leq x_1 < \dots < x_{n+2} \leq b$ such that

$$|f(x_k) - r(x_k)| = \|r - f\| = E_n(f), \quad k = 1, \dots, n + 2.$$

Fix k ($1 \leq k \leq n + 2$). Assume first $f(x_k) - r(x_k) = E_n(f)$. Then, we have

$$f(x_k) - \frac{1}{2}[p(x_k) + q(x_k)] = f(x_k) - r(x_k) = E_n(f) = \|f - p\| \geq f(x_k) - p(x_k),$$

leading to $q(x_k) \leq p(x_k)$. Similarly, we have $p(x_k) \leq q(x_k)$. Thus, $p(x_k) = q(x_k)$. Assume now $f(x_k) - r(x_k) = -E_n(f)$. We have

$$f(x_k) - \frac{1}{2}[p(x_k) + q(x_k)] = f(x_k) - r(x_k) = -E_n(f) = -\|f - p\| \leq f(x_k) - p(x_k),$$

leading to $q(x_k) \geq p(x_k)$. Similarly, $p(x_k) \geq q(x_k)$. Hence, $p(x_k) = q(x_k)$. Therefore, $p(x_k) = q(x_k)$ for all the $n + 2$ distinct points x_1, \dots, x_{n+2} . Hence, $p = q$ in \mathcal{P}_n . \square

1.3 Chebyshev Polynomials

Consider $n \geq 1$ and $f(x) = x^n$ for $x \in [-1, 1]$. By Theorem 1.6, There exists a unique $p_{n-1} \in \mathcal{P}_{n-1}$ such that

$$\|f - p_{n-1}\|_{C[-1,1]} \leq \|f - q_{n-1}\|_{C[-1,1]} \quad \forall q_{n-1} \in \mathcal{P}_{n-1}.$$

Denote for any integer $k \geq 0$

$$\tilde{\mathcal{P}}_k = \{p \in \mathcal{P}_k : \text{the leading coefficient of } p \text{ is } 1\}.$$

Then $\tilde{T}_n := f - p_{n-1} \in \tilde{\mathcal{P}}_n$ is the unique polynomial in $\tilde{\mathcal{P}}_n$ that satisfies

$$\|\tilde{T}_n\|_{C[-1,1]} = \min_{\tilde{p} \in \tilde{\mathcal{P}}_n} \|\tilde{p}\|_{C[-1,1]}. \quad (1.22)$$

By the Chebyshev Alternation Theorem, \tilde{T}_n is characterized by achieving its maximum magnitude at $n + 1$ points in $[-1, 1]$ with alternating signs.

To find \tilde{T}_n , we consider for each $n \geq 0$ ¹

$$T_n(x) = \cos(n \arccos x) \quad \forall x \in [-1, 1]. \quad (1.23)$$

Introducing $x = \cos \theta$ for all $\theta \in [0, \pi]$, we can write

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

Since

$$\cos(n + 1)\theta + \cos(n - 1)\theta = 2 \cos \theta \cos n\theta, \quad (1.24)$$

we obtain

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x) \quad x \in [-1, 1]. \quad (1.25)$$

Clearly, $T_0(x) = 1$ and $T_1(x) = x$. Therefore, by induction, we conclude that T_n is a polynomial of degree n , and for $n \geq 1$ the leading coefficient of T_n is 2^{n-1} .

¹See [8] for more on the derivation of \tilde{T}_n .

Let $\theta_k = k\pi/n \in [0, \pi]$ and $x_k = \cos \theta_k$, $k = 0, \dots, n$. Then,

$$T_n(x_k) = \cos n\theta_k = \cos k\pi = (-1)^k = (-1)^k \|T_n\|_{C[-1,1]}, \quad k = 0, \dots, n.$$

This means that $T_n \in \mathcal{P}_n$ achieves its maximum magnitude at $n + 1$ points in $[-1, 1]$ with alternating signs. Therefore, for $n \geq 1$,

$$\tilde{T}_n(x) = 2^{1-n} T_n(x) = 2^{1-n} \cos(n \arccos x), \quad x \in [-1, 1]. \quad (1.26)$$

We call T_n ($n \geq 0$) the *Chebyshev polynomials of the first kind*. The first few of these polynomials are

$$\begin{aligned} T_0(x) &= 1, \\ T_1(x) &= x, \\ T_2(x) &= 2x^2 - 1, \\ T_3(x) &= 4x^3 - 3x, \\ T_4(x) &= 8x^4 - 8x^2 + 1. \end{aligned}$$

The following theorem summarizes the properties of such polynomials:

Theorem 1.10 (Properties of Chebyshev polynomials of the first kind).

(1) Orthogonality.

$$\int_{-1}^1 \frac{1}{\sqrt{1-x^2}} T_m(x) T_n(x) dx = \begin{cases} 0 & \text{if } m \neq n, \\ \pi & \text{if } m = n = 0, \\ \pi/2 & \text{if } m = n > 0. \end{cases} \quad (1.27)$$

(2) Recurrence.

$$\begin{aligned} T_0(x) &= 1, \\ T_1(x) &= x, \\ T_{n+1}(x) &= 2xT_n(x) - T_{n-1}(x), \quad n = 1, \dots \end{aligned}$$

In particular, each T_n is a polynomial of degree n . Moreover, if n is an even (odd) number, then T_n is an even (odd) function.

(3) Extreme points and zeros.

$$\begin{aligned} \|T_n\|_{C[-1,1]} &= 1, \quad n = 0, \dots, \\ T_n\left(\cos \frac{k\pi}{n}\right) &= (-1)^k = (-1)^k, \quad k = 0, \dots, n, \quad n = 1, \dots, \\ T_n\left(\cos \frac{(2k-1)\pi}{2n}\right) &= 0, \quad k = 1, \dots, n, \quad n = 1, \dots \end{aligned}$$

- (4) Best uniform approximation and least-squares approximation. For each $n \geq 1$, $\tilde{T}_n = 2^{1-n}T_n \in \tilde{\mathcal{P}}_n$ is the unique polynomial in $\tilde{\mathcal{P}}_n$ such that

$$\|\tilde{T}_n\|_{C[-1,1]} = \min_{\tilde{p} \in \tilde{\mathcal{P}}_n} \|\tilde{p}\|_{C[-1,1]} = 2^{1-n},$$

$$\int_{-1}^1 \frac{1}{\sqrt{1-x^2}} [\tilde{T}_n(x)]^2 dx = \min_{\tilde{p}_n \in \tilde{\mathcal{P}}_n} \int_{-1}^1 \frac{1}{\sqrt{1-x^2}} [\tilde{p}_n(x)]^2 dx = 2^{1-2n}\pi.$$

- (5) Differential equation.

$$(1-x^2)T_n''(x) - xT_n'(x) + n^2T_n(x) = 0, \quad n = 0, \dots$$

- (6) The generating function.

$$\frac{1-tx}{1-2tx+t^2} = \sum_{n=0}^{\infty} T_n(x)t^n \quad \forall t \in (-1,1) \quad \forall x \in [-1,1].$$

- (7) Rodrigues' formula.

$$T_n(x) = \frac{(-1)^n}{(2n-1)!!} (1-x^2)^{\frac{1}{2}} \frac{d^n}{dx^n} (1-x^2)^{n-\frac{1}{2}}, \quad n = 0, \dots \quad \square$$

Proof. We only prove Parts (1)–(6). The proof of Part (7) can be found in [9].

- (1) By the change of variable $x = \cos \theta$, we obtain

$$\begin{aligned} \int_{-1}^1 \frac{1}{\sqrt{1-x^2}} T_m(x) T_n(x) dx &= \int_0^\pi \cos m\theta \cos n\theta d\theta \\ &= \frac{1}{2} \int_0^\pi [\cos(m+n)\theta + \cos(m-n)\theta] d\theta. \end{aligned}$$

Since

$$\int_0^\pi \cos k\theta d\theta = 0 \quad \text{for any integer } k \geq 1,$$

simple calculations then lead to (1.27).

- (2) This is proved above, cf. (1.23)–(1.25).

- (3) This follows from the definition (1.23) and direct calculations.

(4) The first minimization property is proved above, cf. (1.22) and (1.26). The second minimization property follows from Theorem 1.26 in Section 1.8 on properties of orthogonal polynomials and (1.27).

- (5) By the definition (1.23), we have

$$T_n'(x) = \frac{n \sin(n \arccos x)}{\sqrt{1-x^2}},$$

$$T_n''(x) = \frac{-n^2 \cos(n \arccos x)}{1 - x^2} + \frac{nx \sin(n \arccos x)}{(1 - x^2)^{3/2}}.$$

These, together with (1.23), imply the desired differential equation.

(6) Let i be the complex unit, i.e., $i^2 = -1$. Denote $\mathcal{R}(z)$ the real part of a complex number z . We have using $x = \cos \theta$ that

$$\sum_{n=0}^{\infty} T_n(x)t^n = \sum_{n=0}^{\infty} \cos(n\theta)t^n = \sum_{n=0}^{\infty} \mathcal{R}(t^n e^{in\theta}) = \mathcal{R}\left(\sum_{n=0}^{\infty} (te^{i\theta})^n\right).$$

Since

$$\sum_{n=0}^{\infty} (te^{i\theta})^n = \frac{1}{1 - te^{i\theta}} = \frac{1}{1 - t \cos \theta - i t \sin \theta} = \frac{1 - t \cos \theta + i t \sin \theta}{(1 - t \cos \theta)^2 + t^2 \sin^2 \theta} = \frac{1 - tx + i t \sin \theta}{1 - 2tx + t^2},$$

we have

$$\sum_{n=0}^{\infty} T_n(x)t^n = \mathcal{R}\left(\sum_{n=0}^{\infty} (te^{i\theta})^n\right) = \frac{1 - tx}{1 - 2tx + t^2},$$

completing the proof. \square

1.4 Uniform Approximation by Trigonometric Polynomials

We denote by $C_{2\pi}$ the set of all real-valued, 2π -periodical functions on \mathbb{R} .

Theorem 1.11 (The Second Weierstrass Approximation Theorem). *Let $f \in C_{2\pi}$. For any $\varepsilon > 0$, there exists $T \in \mathcal{T}$ such that*

$$|f(x) - T(x)| < \varepsilon \quad \forall x \in \mathbb{R}. \quad (1.28)$$

Theorem 1.12 (Best uniform approximations by trigonometric polynomials). *For any $f \in C_{2\pi}$ and any integer $n \geq 0$, there exists a unique best uniform approximation of f in \mathcal{T}_n . Moreover, $T_n \in \mathcal{T}_n$ is the best uniform approximation of f in \mathcal{T}_n if and only if there exist*

1.5 Modulus of Continuity and Jackson's Theorems

Definition 1.13 (Modulus of continuity). *Let $f \in C[a, b]$ and $\delta > 0$. The number*

$$\omega(f; \delta) = \sup_{\substack{x, x' \in [a, b] \\ |x - x'| \leq \delta}} |f(x) - f(x')|$$

is called the modulus of continuity of f over $[a, b]$ with $\delta > 0$.

Clearly, f is uniformly continuous if and only if $\lim_{\delta \rightarrow 0} \omega(f; \delta) = 0$.

Definition 1.14 (Lipschitz continuous functions). Let S be a nonempty subset of \mathbb{R} . A function $f : S \rightarrow \mathbb{R}$ is a Lipschitz function, if there exists $L \geq 0$ such that

$$|f(x) - f(y)| \leq L|x - y| \quad \forall x, y \in S.$$

1.6 Least-Squares Approximation

Definition 1.15 (Weight functions). A weight function on a finite interval $[a, b]$ is a non-negative, integrable function $\rho : (a, b) \rightarrow \mathbb{R}$ that satisfies

$$\int_c^d \rho(x) dx > 0 \quad \text{for any sub-interval } (c, d) \subseteq (a, b).$$

Examples.

- (1) For any $[a, b]$, $\rho(x) \equiv 1$ defines a weight function.
- (2) For $[a, b] = [-1, 1]$, $\rho(x) = 1/\sqrt{1-x^2}$ defines a weight function.
- (3) Let $\rho : [a, b] \rightarrow \mathbb{R}$ satisfy the following: (a) ρ is integrable on $[a, b]$; (b) ρ is continuous in (a, b) ; (c) $\rho(x) \geq 0$ for any $x \in (a, b)$ and ρ has at most finitely many zeros in (a, b) . Then, ρ is a weight function on $[a, b]$.

We assume in the rest of this section that ρ is a weight function on $[a, b]$. We denote by $L_\rho^2(a, b)$ the set of all measurable functions $f : [a, b] \rightarrow \mathbb{R}$ such that

$$\int_a^b \rho(x) [f(x)]^2 dx < \infty.$$

If $\rho(x) = 1$ for all $x \in (a, b)$, then we simply write $L^2(a, b)$ instead of $L_\rho^2(a, b)$. Under the usual addition and scalar multiplication, $L_\rho^2(a, b)$ is a vector space. Clearly, $\mathcal{P} \subset C[a, b] \subseteq L_\rho^2(a, b)$. If $f, g \in L_\rho^2(a, b)$, then

$$\int_a^b \rho(x) |f(x)g(x)| dx \leq \int_a^b \rho(x) \frac{[f(x)]^2 + [g(x)]^2}{2} dx < \infty. \quad (1.29)$$

Thus, ρfg is integrable. In particular, setting $g(x) = 1$, we see that ρf is integrable.

Proposition 1.16 (The Cauchy-Schwarz inequality). We have

$$\left| \int_a^b \rho(x) f(x) g(x) dx \right| \leq \sqrt{\int_a^b \rho(x) [f(x)]^2 dx} \sqrt{\int_a^b \rho(x) [g(x)]^2 dx} \quad \forall f, g \in L_\rho^2(a, b). \quad (1.30)$$

Proof. Consider the non-negative quadratic function $\phi : \mathbb{R} \rightarrow \mathbb{R}$ defined by

$$\begin{aligned}\phi(\lambda) &= \int_a^b \rho(x) [f(x) + \lambda g(x)]^2 dx \\ &= \lambda^2 \int_a^b \rho(x) [g(x)]^2 dx + 2\lambda \int_a^b \rho(x) f(x) g(x) dx + \int_a^b \rho(x) [f(x)]^2 dx \quad \forall \lambda \in \mathbb{R}.\end{aligned}$$

If the leading coefficient of ϕ is 0, then the inequality (1.30) holds true trivially. If it is nonzero, then the discriminant of ϕ is non-positive, i.e.,

$$4 \left[\int_a^b \rho(x) f(x) g(x) dx \right]^2 - 4 \int_a^b \rho(x) [g(x)]^2 dx \int_a^b \rho(x) [f(x)]^2 dx \leq 0.$$

This leads to the inequality in (1.30). □

Definition 1.17 (Least-squares approximation). A least-squares approximation of a given function $f \in L^2_\rho(a, b)$ in \mathcal{P}_n is a polynomial $p_n \in \mathcal{P}_n$ that satisfies

$$\int_a^b \rho(x) [f(x) - p_n(x)]^2 dx \leq \int_a^b \rho(x) [f(x) - q_n(x)]^2 dx \quad \forall q_n \in \mathcal{P}_n. \quad (1.31)$$

Theorem 1.18 (Existence, uniqueness, and characterization of least-squares approximations). Let $f \in L^2_\rho(a, b)$. Let $n \geq 0$ be an integer.

- (1) There exists a unique least-squares approximation of f in \mathcal{P}_n .
- (2) Let $p_n \in \mathcal{P}_n$. Then p_n is the least-squares approximation of f in \mathcal{P}_n if and only if

$$\int_a^b \rho(x) [f(x) - p_n(x)] q_n(x) dx = 0 \quad \forall q_n \in \mathcal{P}_n. \quad (1.32)$$

- (3) If $p_n \in \mathcal{P}_n$ is the least-squares approximation of f in \mathcal{P}_n , then

$$\int_a^b \rho(x) [f(x) - p_n(x)]^2 dx = \int_a^b \rho(x) [f(x)]^2 dx - \int_a^b \rho(x) [p_n(x)]^2 dx. \quad (1.33)$$

To prove this theorem, we introduce the $(n + 1) \times (n + 1)$ matrix

$$G_{n+1} := \left[\int_a^b \rho(x) x^{i+j} dx \right]_{i,j=0}^n. \quad (1.34)$$

In the special case that $[a, b] = [0, 1]$ and $\rho(x) \equiv 1$, this is a Hilbert matrix.

Lemma 1.19. The matrix G_{n+1} is symmetric positive definite.

Proof. Define

$$Q(\xi) = \frac{1}{2} \int_a^b \rho(x) \left(\sum_{i=0}^n \xi_i x^i \right)^2 dx \quad \forall \xi = (\xi_0, \dots, \xi_n) \in \mathbb{R}^{n+1}.$$

Clearly, $Q : \mathbb{R}^{n+1} \rightarrow \mathbb{R}$ is a non-negative quadratic form. If $Q(\xi) = 0$ for some $\xi \in \mathbb{R}^{n+1}$, then, by Lemma 1.20, $\sum_{i=0}^n \xi_i x^i = 0$ for all $x \in [a, b]$. Hence, $\xi = 0$ in \mathbb{R}^{n+1} . Therefore, $Q : \mathbb{R}^{n+1} \rightarrow \mathbb{R}$ is a positive quadratic form.

It is clear that

$$Q(\xi) = \frac{1}{2} \sum_{i,j=0}^n \left(\int_a^b \rho(x) x^{i+j} dx \right) \xi_i \xi_j \quad \forall \xi = (\xi_0, \dots, \xi_n) \in \mathbb{R}^{n+1}.$$

Therefore, G_{n+1} is the matrix associated with the positive quadratic form $Q : \mathbb{R}^{n+1} \rightarrow \mathbb{R}$. Hence, G_{n+1} is symmetric positive definite. \square

Proof of Theorem 1.18. (1) Define

$$F(c) = \int_a^b \rho(x) \left[f(x) - \sum_{i=0}^n c_i x^i \right]^2 dx \quad \forall c = (c_0, \dots, c_n) \in \mathbb{R}^{n+1}. \quad (1.35)$$

Clearly, $F : \mathbb{R}^{n+1} \rightarrow \mathbb{R}$ is a non-negative quadratic function. Direct calculations lead to

$$F(c) = \sum_{i,j=0}^n \left[\int_a^b \rho(x) x^{i+j} dx \right] c_i c_j - 2 \sum_{i=0}^n \left[\int_a^b \rho(x) f(x) x^i dx \right] c_i - \int_a^b \rho(x) [f(x)]^2 dx.$$

Thus, the Hessian matrix of F is

$$\left[\frac{\partial^2 F}{\partial c_i \partial c_j} \right]_{i,j=0}^n = 2 \left[\int_a^b \rho(x) x^{i+j} dx \right]_{i,j=0}^n = 2G_{n+1},$$

where G_{n+1} is the matrix defined in (1.34). By Lemma 1.19, the Hessian matrix of F is therefore symmetric positive definite. Consequently, $F : \mathbb{R}^{n+1} \rightarrow \mathbb{R}$ is a convex quadratic form. Since it is non-negative, it admits a unique minimizer $\hat{c} = (\hat{c}_0, \dots, \hat{c}_n) \in \mathbb{R}^{n+1}$. Define

$$p_n(x) = \sum_{i=0}^n \hat{c}_i x^i \quad \forall x \in \mathbb{R}. \quad (1.36)$$

Then, $p_n \in \mathcal{P}_n$ is the unique polynomial in \mathcal{P}_n that satisfies (1.31).

(2) From Part (1), the unique least-squares approximation $p_n \in \mathcal{P}_n$ is given by (1.36), where $\hat{c} = (\hat{c}_0, \dots, \hat{c}_n) \in \mathbb{R}^{n+1}$ is the unique minimizer of the convex quadratic function $F : \mathbb{R}^{n+1} \rightarrow \mathbb{R}$. Clearly, \hat{c} is the unique critical point of F determined by

$$\partial_j F(\hat{c}) = 0, \quad j = 0, \dots, n. \quad (1.37)$$

By the definition of $F : \mathbb{R}^{n+1} \rightarrow \mathbb{R}$ (cf. (1.35)) and the Chain Rule, we obtain for any $c = (c_0, \dots, c_n) \in \mathbb{R}^{n+1}$ that

$$\partial_j F(c) = 2 \int_a^b \rho(x) \left[f(x) - \sum_{i=0}^n c_i x^i \right] (-x^j) dx, \quad j = 0, \dots, n. \quad (1.38)$$

Therefore, by the definition of p_n in (1.36), the system of equations (1.37) is equivalent to

$$\int_a^b \rho(x) [f(x) - p_n(x)] x^j dx, \quad j = 0, \dots, n,$$

which in turn is equivalent to (1.32), since $\mathcal{P}_n = \text{Span}\{1, x, \dots, x^n\}$.

(3) Suppose $p_n \in \mathcal{P}_n$ is the least-squares approximation of f . By (1.32) with $q_n = p_n$, we have

$$\int_a^b \rho(x) [f(x) - p_n(x)] p_n(x) dx = 0.$$

Therefore,

$$\begin{aligned} \int_a^b \rho(x) [f(x)]^2 dx &= \int_a^b \rho(x) [f(x) - p_n(x) + p_n(x)]^2 dx \\ &= \int_a^b \rho(x) [f(x) - p_n(x)]^2 dx + \int_a^b \rho(x) [p_n(x)]^2 dx \\ &\quad + 2 \int_a^b \rho(x) [f(x) - p_n(x)] p_n(x) dx \\ &= \int_a^b \rho(x) [f(x) - p_n(x)]^2 dx + \int_a^b \rho(x) [p_n(x)]^2 dx. \end{aligned}$$

This implies (1.33). □

From the proof, cf. (1.37) and (1.38), we see that this least-squares approximation $p_n(x) = \sum_{i=0}^n \hat{c}_i x^i$ can be obtained by solving the system of linear equations

$$\sum_{i=0}^n \left[\int_a^b \rho(x) x^{i+j} dx \right] \hat{c}_i = \int_a^b \rho(x) f(x) x^j dx \quad j = 0, \dots, n.$$

The coefficient matrix of this system of linear equations is exactly G_{n+1} , defined in (1.34).

1.7 Orthogonal Polynomials

Fix a weight function ρ on $[a, b]$. For convenience, we denote

$$\langle f, g \rangle = \int_a^b \rho(x) f(x) g(x) dx \quad \forall f, g \in L_\rho^2(a, b).$$

By (1.29), $\langle f, g \rangle$ is finite for any $f, g \in L^2_\rho(a, b)$. The mapping $\langle \cdot, \cdot \rangle : L^2_\rho(a, b) \times L^2_\rho(a, b) \rightarrow \mathbb{R}$ clearly satisfies the following properties:

(1) Symmetry.

$$\langle f, g \rangle = \langle g, f \rangle \quad \forall f, g \in L^2_\rho(a, b);$$

(2) Bilinearity.

$$\langle \alpha f + \beta g, h \rangle = \alpha \langle f, h \rangle + \beta \langle g, h \rangle \quad \forall f, g, h \in L^2_\rho(a, b) \quad \forall \alpha, \beta \in \mathbb{R};$$

(3) Non-negativity.

$$\langle f, f \rangle \geq 0 \quad \forall f \in L^2_\rho(a, b).$$

Lemma 1.20. *Let $f \in C[a, b]$. If*

$$\int_a^b \rho(x)[f(x)]^2 dx = 0,$$

then $f(x) = 0$ for all $x \in [a, b]$.

Proof. Suppose there existed a point $x_0 \in [a, b]$ such that $f(x_0) \neq 0$. By the continuity of f , there would have existed a sub-interval $[c, d] \subseteq [a, b]$ with $c < d$ such that $|f(x)| \geq \varepsilon_0$ for all $x \in [c, d]$ for some constant $\varepsilon_0 > 0$. Now, by the definition of a weight function, $\int_c^d \rho(x) dx > 0$. Therefore,

$$\int_a^b \rho(x)[f(x)]^2 dx \geq \int_c^d \rho(x)[f(x)]^2 dx \geq \varepsilon_0^2 \int_c^d \rho(x) dx > 0.$$

This contradicts the assumption. □

Consider the subspace $C[a, b]$ of $L^2_\rho(a, b)$. The mapping $\langle \cdot, \cdot \rangle : C[a, b] \times C[a, b] \rightarrow \mathbb{R}$ is in fact an inner product of $C[a, b]$, i.e., it is symmetric, bilinear, and positive. Here, the positivity means

$$\begin{aligned} \langle f, f \rangle &\geq 0 && \forall f \in C[a, b]. \\ \langle f, f \rangle &= 0 && \text{if and only if } f = 0 \text{ in } C[a, b]. \end{aligned}$$

The “only if” part in the last equation follows from Lemma 1.20. The norm of $C[a, b]$ induced by this inner product is

$$\|f\|_{L^2_\rho(a, b)} = \sqrt{\langle f, f \rangle} = \sqrt{\int_a^b \rho(x)[f(x)]^2 dx} \quad (1.39)$$

for any $f \in C[a, b]$. This is called the $L^2_\rho(a, b)$ -norm. We will also use (1.39) for any $f \in L^2_\rho(a, b)$. It is often written as $\|\cdot\|$ instead of $\|\cdot\|_{L^2_\rho(a, b)}$, when no confusion arises.

Definition 1.21 (Orthogonal polynomials). A sequence of polynomials Q_n ($n = 0, \dots$) are called orthogonal polynomials in $L^2_\rho(a, b)$, if the following hold true:

- (1) For each integer $n \geq 0$, Q_n is a polynomial of degree exactly n ;
- (2) $\langle Q_m, Q_n \rangle = 0$ whenever $m \neq n$.

A sequence of orthogonal polynomials Q_n ($n = 0, \dots$) in $L^2_\rho(a, b)$ are called orthonormal polynomials in $L^2_\rho(a, b)$, if they satisfy

- (3) $\langle Q_n, Q_n \rangle = 1$ for all $n \geq 0$.

A set of finitely many polynomials are called orthogonal (or orthonormal) in $L^2_\rho(a, b)$ if it is a subset of a sequence of orthogonal (or orthonormal) polynomials.

If Q_n ($n = 0, \dots$) are orthogonal polynomials, then $\langle Q_n, Q_n \rangle \neq 0$ for each $n \geq 0$. Moreover, the polynomials $Q_n/\sqrt{\langle Q_n, Q_n \rangle}$ ($n = 0, \dots$) are orthonormal. In general, orthonormal polynomials Q_n ($n = 0, \dots$) are characterized by

$$\langle Q_i, Q_j \rangle = \delta_{i,j} \quad \forall i, j \geq 0, \quad (1.40)$$

where δ_{ij} is defined to be 1 if $i = j$ and 0 if $i \neq j$.

The Chebyshev polynomials T_n ($n = 0, \dots$) are orthogonal polynomials in $L^2_\rho(-1, 1)$ with $\rho(x) = 1/\sqrt{1-x^2}$. Another important class of orthogonal polynomials are Legendre polynomials P_n ($n = 0, \dots$) in $L^2(-1, 1)$ that will be discussed in Section 1.9.

Lemma 1.22. If Q_0, \dots, Q_n be orthogonal polynomials in $L^2_\rho(a, b)$, then

- (1) The $n + 1$ polynomials Q_0, \dots, Q_n are linearly independent in $L^2_\rho(a, b)$;
- (2) $\mathcal{P}_n = \text{Span}\{Q_0, \dots, Q_n\}$. Moreover,

$$p_n = \sum_{k=0}^n \frac{\langle p_n, Q_k \rangle}{\langle Q_k, Q_k \rangle} Q_k \quad \forall p_n \in \mathcal{P}_n. \quad (1.41)$$

Proof. (1) Suppose there exist $\alpha_0, \dots, \alpha_n \in \mathbb{R}$ such that $\sum_{k=0}^n \alpha_k Q_k = 0$. Multiplying both sides of this equation by ρQ_j for an arbitrary but fixed j with $0 \leq j \leq n$ and then integrate over $[a, b]$, we obtain by (1.40) that $\alpha_j \langle Q_j, Q_j \rangle = 0$, implying $\alpha_j = 0$, since $\langle Q_j, Q_j \rangle \neq 0$ by Lemma 1.20. Hence, Q_0, \dots, Q_n are linearly independent.

(2) Since the dimension of \mathcal{P}_n is $n + 1$, the $n + 1$ linearly independent polynomials Q_0, \dots, Q_n in \mathcal{P}_n form a basis of \mathcal{P}_n . Hence, they span \mathcal{P}_n .

Fix $p_n \in \mathcal{P}_n = \text{Span}\{Q_0, \dots, Q_n\}$. There exist constants $c_i \in \mathbb{R}$ ($i = 0, \dots, n$) such that $p_n = \sum_{k=0}^n c_k Q_k$. Multiplying both sides of this equation by ρQ_j for an arbitrary but fixed j with $0 \leq j \leq n$ and then integrate over $[a, b]$, we obtain by (1.40) that $\langle p_n, Q_j \rangle = c_j \langle Q_j, Q_j \rangle$. This implies (1.41). \square

Theorem 1.23 (Characterization of least-squares approximations using orthonormal polynomials). Let Q_0, \dots, Q_n be orthonormal polynomials in $L^2_\rho(a, b)$. Then the least-squares approximation of a given $f \in L^2_\rho(a, b)$ in \mathcal{P}_n is given by

$$p_n = \sum_{k=0}^n \langle f, Q_k \rangle Q_k. \quad (1.42)$$

Moreover,

$$\|f - q_n\|^2 = \|f\|^2 - \sum_{k=0}^n \langle f, Q_k \rangle^2. \quad (1.43)$$

Proof. Clearly, the polynomial p_n defined in (1.42) is in \mathcal{P}_n . Fix an arbitrary integer j with $0 \leq j \leq n$. Since Q_0, \dots, Q_n are orthonormal, we have

$$\langle f - p_n, Q_j \rangle = \langle f, Q_j \rangle - \sum_{k=0}^n \langle f, Q_k \rangle \langle Q_k, Q_j \rangle = \langle f, Q_j \rangle - \langle f, Q_j \rangle = 0.$$

Hence, by Lemma 1.22,

$$\langle f - p_n, q_n \rangle = 0 \quad \forall q_n \in \mathcal{P}_n.$$

This is exactly (1.32). Therefore, by Theorem 1.18, $p_n \in \mathcal{P}_n$ is the unique least-squares approximation of f in \mathcal{P}_n .

It follows from (1.42), the symmetry and bilinearity of $\langle \cdot, \cdot \rangle$, and the fact that Q_0, \dots, Q_n are orthonormal that

$$\langle p_n, p_n \rangle = \left\langle \sum_{i=0}^n \langle f, Q_i \rangle Q_i, \sum_{i=0}^n \langle f, Q_i \rangle Q_i \right\rangle = \sum_{i,j=0}^n \langle f, Q_i \rangle \langle f, Q_j \rangle \langle Q_i, Q_j \rangle = \sum_{i=0}^n \langle f, Q_i \rangle^2.$$

This, together with Part (3) of Theorem 1.18, implies (1.43). \square

Theorem 1.24 (The Gram-Schmidt orthogonalization). Consider $L^2_\rho(a, b)$. Let $f_0(x) = 1$ and $f_k(x) = x^k$ for any integer $k \geq 1$. Define

$$g_0(x) = \frac{1}{\sqrt{\langle f_0, f_0 \rangle}} f_0(x), \quad (1.44)$$

$$\begin{cases} \hat{g}_k(x) = f_k(x) - \sum_{j=1}^{k-1} \langle f_k, g_j \rangle g_j(x), \\ g_k(x) = \frac{1}{\sqrt{\langle \hat{g}_k, \hat{g}_k \rangle}} \hat{g}_k(x), \end{cases} \quad k = 2, \dots \quad (1.45)$$

Then g_k ($k = 0, \dots$) are orthonormal polynomials.

Proof. Clearly, g_0 and g_1 are polynomial of degrees 0 and 1, respectively. Moreover, direct calculations using (1.44) and (1.45) lead to $\langle g_0, g_0 \rangle = \langle g_1, g_1 \rangle = 1$ and $\langle g_0, g_1 \rangle = 0$. Let $k \geq 1$ be an integer. Assume that g_i is a polynomial of degree i for each i with $0 \leq i \leq k$ and that

$$\langle g_i, g_j \rangle = \delta_{ij}, \quad 0 \leq i, j \leq k. \quad (1.46)$$

Since f_{k+1} is a polynomial of degree $k + 1$, we see from (1.45) with k replaced by $k + 1$ that g_{k+1} is a polynomial of degree $k + 1$. Clearly, $\langle g_{k+1}, g_{k+1} \rangle = 1$. Moreover, for each i with $0 \leq i \leq k$, we have by (1.45) and (1.46) that

$$\langle \hat{g}_{k+1}, g_i \rangle = \langle f_{k+1}, g_i \rangle - \sum_{j=0}^k \langle f_{k+1}, g_j \rangle \langle g_j, g_i \rangle = \langle f_{k+1}, g_i \rangle - \langle f_{k+1}, g_i \rangle = 0.$$

Since g_{k+1} and \hat{g}_{k+1} differ by a nonzero constant multiplier, we have

$$\langle g_{k+1}, g_i \rangle = 0, \quad i = 0, \dots, k.$$

Therefore, (1.46) is true with k replaced by $k + 1$. Consequently, g_k ($k = 0, \dots$) are orthonormal polynomials. \square

Example. Consider $L^2[-1, 1]$.

$$\begin{aligned} g_0(x) &= \frac{1}{\sqrt{\langle f_0, f_0 \rangle}} f_0(x) = \frac{1}{\sqrt{\int_{-1}^1 1 \cdot 1 dx}} = \frac{1}{\sqrt{2}}. \\ \hat{g}_1(x) &= f_1(x) - \langle f_1, g_0 \rangle g_0(x) = x - \left(\int_{-1}^1 \frac{x}{\sqrt{2}} dx \right) \frac{1}{\sqrt{2}} = x, \\ g_1(x) &= \frac{1}{\sqrt{\langle \hat{g}_1, \hat{g}_1 \rangle}} \hat{g}_1(x) = \frac{x}{\sqrt{\int_{-1}^1 x^2 dx}} = \sqrt{\frac{3}{2}} x. \\ \hat{g}_2(x) &= f_2(x) - \langle f_2, g_0 \rangle g_0(x) - \langle f_2, g_1 \rangle g_1(x) \\ &= x^2 - \left(\int_{-1}^1 \frac{x^2}{\sqrt{2}} dx \right) \frac{1}{\sqrt{2}} - \left(\int_{-1}^1 x^2 \sqrt{\frac{3}{2}} x dx \right) \sqrt{\frac{3}{2}} x = x^2 - \frac{1}{3}, \\ g_2(x) &= \frac{1}{\sqrt{\langle \hat{g}_2, \hat{g}_2 \rangle}} \hat{g}_2(x) = \frac{1}{\sqrt{\int_{-1}^1 (x^2 - 1/3)^2 dx}} \left(x^2 - \frac{1}{3} \right) = \frac{\sqrt{15}}{4} \left(x^2 - \frac{1}{3} \right). \end{aligned}$$

Therefore, g_0, g_1, g_2 are orthonormal polynomials in $L^2[-1, 1]$.

1.8 More Properties of Orthogonal Polynomials

Theorem 1.25 (The three-term recurrence). Consider $L^2_\rho(a, b)$.

(1) Define

$$\begin{aligned} \tilde{Q}_0(x) &= 1, \\ \tilde{Q}_1(x) &= x - a_1, \\ \tilde{Q}_n(x) &= (x - a_n) \tilde{Q}_{n-1}(x) - b_n \tilde{Q}_{n-2}(x), \quad n = 2, \dots, \end{aligned}$$

$$a_n = \frac{\langle x\tilde{Q}_{n-1}, \tilde{Q}_{n-1} \rangle}{\langle \tilde{Q}_{n-1}, \tilde{Q}_{n-1} \rangle}, \quad n = 1, \dots,$$

$$b_n = \frac{\langle \tilde{Q}_{n-1}, \tilde{Q}_{n-1} \rangle}{\langle \tilde{Q}_{n-2}, \tilde{Q}_{n-2} \rangle}, \quad n = 2, \dots$$

Then \tilde{Q}_n ($n = 0, \dots$) are the orthogonal polynomials with each $\tilde{Q}_n \in \tilde{\mathcal{P}}_n$.

(2) Let Q_n ($n = 0, \dots$) be orthogonal polynomials. Then

$$Q_n(x) = (A_n x + B_n)Q_{n-1}(x) - C_n Q_{n-2}(x) \quad n = 2, \dots \quad (1.47)$$

where

$$A_n = \frac{\alpha_n}{\alpha_{n-1}}, \quad B_n = -\frac{\alpha_n}{\alpha_{n-1}} \frac{\langle xQ_{n-1}, Q_{n-1} \rangle}{\langle Q_{n-1}, Q_{n-1} \rangle}, \quad C_n = \frac{\alpha_n \alpha_{n-2}}{\alpha_{n-1}^2} \frac{\langle Q_{n-1}, Q_{n-1} \rangle}{\langle Q_{n-2}, Q_{n-2} \rangle}.$$

Proof. (1) Clearly, $\tilde{Q}_0 \in \tilde{\mathcal{P}}_0$ and $\tilde{Q}_1 \in \tilde{\mathcal{P}}_1$. Since

$$a_1 = \frac{\langle x\tilde{Q}_0, \tilde{Q}_0 \rangle}{\langle \tilde{Q}_0, \tilde{Q}_0 \rangle} = \frac{\langle x, 1 \rangle}{\langle 1, 1 \rangle},$$

we have

$$\langle \tilde{Q}_0, \tilde{Q}_1 \rangle = \langle 1, x - a_1 \rangle = \langle 1, x \rangle - \langle 1, 1 \rangle a_1 = 0.$$

Let $n \geq 2$. Assume $\tilde{Q}_k \in \tilde{\mathcal{P}}_k$ for all $k \in \{0, \dots, n-1\}$ and

$$\langle \tilde{Q}_j, \tilde{Q}_k \rangle = 0, \quad \text{if } 0 \leq j, k \leq n-1 \text{ and } j \neq k. \quad (1.48)$$

Clearly, $\tilde{Q}_n \in \tilde{\mathcal{P}}_n$. We have by the definition of \tilde{Q}_n and (1.48) that

$$\begin{aligned} \langle \tilde{Q}_n, \tilde{Q}_{n-1} \rangle &= \langle (x - a_n)\tilde{Q}_{n-1} - b_n\tilde{Q}_{n-2}, \tilde{Q}_{n-1} \rangle \\ &= \langle x\tilde{Q}_{n-1}, \tilde{Q}_{n-1} \rangle - a_n \langle \tilde{Q}_{n-1}, \tilde{Q}_{n-1} \rangle - b_n \langle \tilde{Q}_{n-2}, \tilde{Q}_{n-1} \rangle \\ &= \langle x\tilde{Q}_{n-1}, \tilde{Q}_{n-1} \rangle - \frac{\langle x\tilde{Q}_{n-1}, \tilde{Q}_{n-1} \rangle}{\langle \tilde{Q}_{n-1}, \tilde{Q}_{n-1} \rangle} \langle \tilde{Q}_{n-1}, \tilde{Q}_{n-1} \rangle \\ &= 0, \\ \langle \tilde{Q}_n, \tilde{Q}_{n-2} \rangle &= \langle (x - a_n)\tilde{Q}_{n-1} - b_n\tilde{Q}_{n-2}, \tilde{Q}_{n-2} \rangle \\ &= \langle x\tilde{Q}_{n-1}, \tilde{Q}_{n-2} \rangle - a_n \langle \tilde{Q}_{n-1}, \tilde{Q}_{n-2} \rangle - b_n \langle \tilde{Q}_{n-2}, \tilde{Q}_{n-2} \rangle \\ &= \langle x\tilde{Q}_{n-1}, \tilde{Q}_{n-2} \rangle - \frac{\langle \tilde{Q}_{n-1}, \tilde{Q}_{n-1} \rangle}{\langle \tilde{Q}_{n-2}, \tilde{Q}_{n-2} \rangle} \langle \tilde{Q}_{n-2}, \tilde{Q}_{n-2} \rangle \\ &= \langle \tilde{Q}_{n-1}, x\tilde{Q}_{n-2} \rangle - \langle \tilde{Q}_{n-1}, \tilde{Q}_{n-1} \rangle \end{aligned}$$

$$\begin{aligned}
&= \langle \tilde{Q}_{n-1}, x\tilde{Q}_{n-2} - \tilde{Q}_{n-1} \rangle \\
&= 0,
\end{aligned}$$

where the last equation follows from that fact that $x\tilde{Q}_{n-2} - \tilde{Q}_{n-1} \in \mathcal{P}_{n-2}$. Let $0 \leq k \leq n-3$ with $n \geq 3$. We have again by the definition of \tilde{Q}_n and (1.48) that

$$\begin{aligned}
\langle \tilde{Q}_n, \tilde{Q}_k \rangle &= \langle (x - a_n)\tilde{Q}_{n-1} - b_n\tilde{Q}_{n-2}, \tilde{Q}_k \rangle \\
&= \langle x\tilde{Q}_{n-1}, \tilde{Q}_k \rangle - a_n \langle \tilde{Q}_{n-1}, \tilde{Q}_k \rangle - b_n \langle \tilde{Q}_{n-2}, \tilde{Q}_k \rangle \\
&= \langle \tilde{Q}_{n-1}, x\tilde{Q}_k \rangle \\
&= 0,
\end{aligned}$$

since $x\tilde{Q}_k \in \mathcal{P}_{k+1} \subseteq \mathcal{P}_{n-2}$. Therefore, (1.48) holds true with $n-1$ replaced by n . Part (1) is thus proved.

(2) Note that $\tilde{Q}_n := (1/\alpha_n)Q_n \in \tilde{\mathcal{P}}_n$. Thus, it follows from Part (1) and the definition of all a_n, b_n , and A_n, B_n, C_n that

$$\begin{aligned}
Q_n(x) &= \alpha_n \tilde{Q}_n(x) \\
&= \alpha_n [(x - a_n)\tilde{Q}_{n-1}(x) - b_n\tilde{Q}_{n-2}(x)] \\
&= \alpha_n \left[(x - a_n) \frac{1}{\alpha_{n-1}} Q_{n-1}(x) - b_n \frac{1}{\alpha_{n-2}} Q_{n-2}(x) \right] \\
&= \left(\frac{\alpha_n}{\alpha_{n-1}} x - \frac{\alpha_n a_n}{\alpha_{n-1}} \right) Q_{n-1}(x) - \frac{\alpha_n b_n}{\alpha_{n-2}} Q_{n-2}(x) \\
&= (A_n x + B_n) Q_{n-1}(x) - C_n Q_{n-2}(x) \quad \forall n \geq 2.
\end{aligned}$$

This is (1.47). □

Example. Consider $L^2[-1, 1]$.

$$\begin{aligned}
\tilde{Q}_0(x) &= 1. \\
a_1 &= \frac{\langle x\tilde{Q}_0, \tilde{Q}_0 \rangle}{\langle \tilde{Q}_0, \tilde{Q}_0 \rangle} = 0, \\
\tilde{Q}_1(x) &= x - a_1 = x. \\
a_2 &= \frac{\langle x\tilde{Q}_1, \tilde{Q}_1 \rangle}{\langle \tilde{Q}_1, \tilde{Q}_1 \rangle} = \frac{\int_{-1}^1 x^3 dx}{\int_{-1}^1 x^2 dx} = 0, \\
b_2 &= \frac{\langle \tilde{Q}_1, \tilde{Q}_1 \rangle}{\langle \tilde{Q}_0, \tilde{Q}_0 \rangle} = \frac{\int_{-1}^1 x^2 dx}{\int_{-1}^1 1 dx} = \frac{1}{3}, \\
\tilde{Q}_2(x) &= (x - a_2)\tilde{Q}_1(x) - b_2\tilde{Q}_0(x) = x^2 - \frac{1}{3}.
\end{aligned}$$

Thus, $\tilde{Q}_0, \tilde{Q}_1, \tilde{Q}_2$ are the orthogonal polynomials in $L^2[-1, 1]$ that have leading coefficient 1. Each of them differs by a constant multiplier from the corresponding one obtained in the previous example using the Gram-Schmidt orthogonalization.

Theorem 1.26 (Minimization). *Let Q_n ($n = 0, \dots$) be orthogonal polynomials in $L^2_\rho(a, b)$. Suppose $n \geq 1$ and $Q_n \in \tilde{\mathcal{P}}_n$, then Q_n is the unique polynomial in $\tilde{\mathcal{P}}_n$ that satisfies*

$$\|Q_n\| = \min_{q_n \in \tilde{\mathcal{P}}_n} \|q_n\|.$$

Proof. Since $Q_n \in \tilde{\mathcal{P}}_n$, we have $Q_n(x) = x^n - q_{n-1}(x)$ for some $q_{n-1} \in \mathcal{P}_{n-1}$. Thus, by the orthogonality,

$$0 = \langle Q_n, q \rangle = \langle x^n - q_{n-1}, q \rangle \quad \forall q \in \mathcal{P}_{n-1}.$$

Theorem 1.18 then implies that q_{n-1} is the unique least-squares approximation of x^n in \mathcal{P}_{n-1} . This is equivalent to the assertion of the theorem. \square

Colloary 1.27 (Uniqueness of orthogonal polynomials). *If $\{P_n\}_{n=0}^\infty$ and $\{Q_n\}_{n=0}^\infty$ are two systems of orthogonal polynomials in $L^2_\rho(a, b)$, then, for each $n \geq 0$, there exists $c_n \in \mathbb{R}$ with $c_n \neq 0$ such that $P_n = c_n Q_n$.*

Proof. Let α_n and β_n be the leading coefficients of P_n and Q_n , respectively. Then, by Theorem 1.26, $(1/\alpha_n)P_n = (1/\beta_n)Q_n$ for each $n \geq 1$. This is clearly true also for $n = 0$. Thus, $P_n = c_n Q_n$ with $c_n = \alpha_n/\beta_n$ for each $n \geq 0$. \square

Theorem 1.28 (Zeros of orthogonal polynomials). *Let Q_n ($n = 0, \dots$) be orthogonal polynomials in $L^2_\rho(a, b)$. Then, for each $n \geq 1$, Q_n has exactly n simple roots in (a, b) .*

Proof. Fix an integer $n \geq 1$. By the orthogonality,

$$\int_a^b \rho(x) Q_n(x) dx = \langle Q_n, 1 \rangle = 0.$$

Hence, Q_n changes its sign in (a, b) at least once. If $n = 1$, this implies that Q_1 has exactly one root. Consider $n \geq 2$. Suppose Q_n changes its sign in (a, b) only k times with $1 \leq k \leq n - 1$ at x_1, \dots, x_k with $a < x_1 < \dots < x_k < b$. Define

$$p(x) = (x - x_1) \cdots (x - x_k).$$

Clearly, $p \in \mathcal{P}_k \subseteq \mathcal{P}_{n-1}$. Moreover, both Q_n and p change their signs only at x_1, \dots, x_k . Thus,

$$\langle Q_n, p \rangle = \int_a^b \rho(x) Q_n(x) p(x) dx \neq 0.$$

This contradicts the fact that $\langle Q_n, q \rangle = 0$ for any $q \in \mathcal{P}_{n-1}$. Hence, $k \geq n$. But $Q_n \in \mathcal{P}_n$ can have at most n roots. Thus, $k = n$, and Q_n has exactly n simple roots in (a, b) . \square

Let Q_n ($n = 0, \dots$) be orthonormal polynomials in $L^2_\rho(a, b)$ and define for each $n \geq 0$

$$K_n(x, t) = \sum_{k=0}^n Q_k(x)Q_k(t). \quad (1.49)$$

We call $K_n : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ the *Dirichlet kernel* associated with first $n + 1$ orthonormal polynomials Q_0, \dots, Q_n in $L^2_\rho(a, b)$. We have for any $f \in L^2_\rho(a, b)$ that

$$\begin{aligned} \sum_{k=0}^n \langle f, Q_k \rangle Q_k(x) &= \sum_{k=0}^n \int_a^b \rho(t) f(t) Q_k(t) dt Q_k(x) \\ &= \sum_{k=0}^n \int_a^b \rho(t) f(t) Q_k(t) Q_k(x) dt \\ &= \int_a^b \rho(t) K_n(x, t) f(t) dt \\ &= \langle K_n(x, \cdot), f(\cdot) \rangle. \end{aligned}$$

It then follows from (1.42) in Theorem 1.23 that the least-squares approximation of $f \in L^2_\rho(a, b)$ in \mathcal{P}_n is given by

$$p_n(x) = \langle K_n(x, \cdot), f(\cdot) \rangle.$$

Theorem 1.29 (The Christoffel-Darboux identity). *Let Q_n ($n = 0, \dots$) be orthonormal polynomials in $L^2_\rho[a, b]$ and $K_n : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ the Dirichlet kernel associated with Q_0, \dots, Q_n for each $n \geq 0$. Then,*

$$K_n(x, t) = \frac{\alpha_n}{\alpha_{n+1}} \frac{Q_{n+1}(x)Q_n(t) - Q_{n+1}(t)Q_n(x)}{x - t} \quad \forall x, t \in \mathbb{R} \text{ with } x \neq t, \quad (1.50)$$

where α_k is the leading coefficient of Q_k ($k = 0, \dots$).

The formula (1.50) is called the *Christoffel-Darboux identity*.

Proof of Theorem 1.29. By Theorem 1.25, the orthonormal polynomials Q_n ($n = 0, \dots$) satisfy (1.47) with $A_n = \alpha_n/\alpha_{n-1}$ and $C_n = A_n/A_{n-1}$, where α_n is the leading coefficient of Q_n . Therefore, for a fixed $n \geq 1$,

$$\begin{aligned} &A_{k+1}^{-1} [Q_{k+1}(x)Q_k(t) - Q_{k+1}(t)Q_k(x)] \\ &= A_{k+1}^{-1} [(A_{k+1}x + B_{k+1})Q_k(x) - C_{k+1}Q_{k-1}(x)] Q_k(t) \\ &\quad - A_{k+1}^{-1} [(A_{k+1}t + B_{k+1})Q_k(t) - C_{k+1}Q_{k-1}(t)] Q_k(x) \\ &= (x - t)Q_k(x)Q_k(t) + A_k^{-1} [Q_k(x)Q_{k-1}(t) - Q_k(t)Q_{k-1}(x)], \quad k = 1, \dots, n. \end{aligned}$$

Summing over these, we obtain that

$$A_{n+1}^{-1} [Q_{n+1}(x)Q_n(t) - Q_{n+1}(t)Q_n(x)]$$

$$\begin{aligned}
&= (x-t) \sum_{k=1}^n Q_k(x)Q_k(t) + A_1^{-1} [Q_1(x)Q_0(t) - Q_1(t)Q_0(x)] \\
&= (x-t) \sum_{k=1}^n Q_k(x)Q_k(t) + (x-t)Q_0(x)Q_0(t),
\end{aligned}$$

since $Q_0(x) = Q_0(t) = \alpha_0$. This leads to (1.50) for $n \geq 2$. For the case $n = 0$ or 1 , one can directly verify that (1.50) is true. \square

1.9 Legendre Polynomials

The Legendre polynomials $P_n \in \mathcal{P}_n$ ($n = 0, \dots$) are the unique orthogonal polynomials in $L^2(-1, 1)$ that are normalized by

$$P_n(1) = 1 \quad \forall n \geq 0.$$

A convenient way to define these polynomials is to use *Rodrigues' formula*:

$$P_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} [(x^2 - 1)^n], \quad n = 0, \dots$$

Clearly, for each integer $n \geq 0$, P_n is a polynomial of degree exactly n . Moreover, if n is even (odd), then P_n is an even (odd) polynomial, i.e., it only consists of even (odd) powers of x . The first few of these polynomials are

$$\begin{aligned}
P_0(x) &= 1, \\
P_1(x) &= x, \\
P_2(x) &= \frac{3}{2}x^2 - \frac{1}{2}, \\
P_3(x) &= \frac{5}{2}x^3 - \frac{3}{2}x, \\
P_4(x) &= \frac{35}{8}x^4 - \frac{15}{4}x^2 + \frac{3}{8}.
\end{aligned}$$

We summarize in the following theorem some of the important properties of Legendre polynomials:

Theorem 1.30 (Properties of Legendre polynomials).

(1) Orthogonality.

$$\int_{-1}^1 P_m(x)P_n(x) dx = \begin{cases} 0 & \text{if } m \neq n, \\ 2/(2n+1) & \text{if } m = n. \end{cases} \quad (1.51)$$

(2) Recurrence.

$$\begin{aligned} P_0(x) &= 1, \\ P_1(x) &= x, \\ (n+1)P_{n+1}(x) - (2n+1)xP_n(x) + nP_{n-1}(x) &= 0, \quad n = 1, \dots \end{aligned} \quad (1.52)$$

(3) Zeros. For each $n \geq 1$, P_n has n simple roots in $(-1, 1)$.

(4) The least-squares approximation. For each $n \geq 1$, $\tilde{P}_n := [2^n(n!)^2/(2n)!]P_n \in \mathcal{P}_n$ is the unique polynomial in $\tilde{\mathcal{P}}_n$ such that

$$\|\tilde{P}_n\|_{L^2(-1,1)} = \frac{2^n(n!)^2}{(2n)!} \sqrt{\frac{2}{2n+1}} = \min_{\tilde{p}_n \in \tilde{\mathcal{P}}_n} \|\tilde{p}_n\|_{L^2(-1,1)}.$$

(5) Normalization. For each $n \geq 0$,

$$P_n(1) = (-1)^n P_n(-1) = 1.$$

(6) Differential equation. For all $n \geq 0$,

$$(1-x^2)P_n''(x) - 2xP_n'(x) + n(n+1)P_n(x) = 0. \quad (1.53)$$

(7) The generating function. For any $x \in [-1, 1]$, the values $P_n(x)$ ($n = 0, \dots$) are the coefficients of the Maclaurin series

$$\frac{1}{\sqrt{1-2tx+t^2}} = \sum_{n=0}^{\infty} P_n(x)t^n \quad \forall t \in (-1, 1).$$

Proof. We only prove Parts (1)–(6). The proof of Part (7) can be found in [9].

(1) Let m, n be non-negative integers. If one of them is 0, then it is clear that (1.51) holds true. Assume $1 \leq m \leq n$. Let $\phi_k(x) = (x^2 - 1)^k$ for any integer $k \geq 0$. Then, $P_k = [1/(2^k k!)]\phi_k^{(k)}$. By integration by parts, we thus have

$$\begin{aligned} \int_{-1}^1 \phi_m^{(m)}(x)\phi_n^{(n)}(x) dx &= \phi_m^{(m)}(x)\phi_n^{(n-1)}(x)\Big|_{x=-1}^{x=1} - \int_{-1}^1 \phi_m^{(m+1)}(x)\phi_n^{(n-1)}(x) dx \\ &= - \int_{-1}^1 \phi_m^{(m+1)}(x)\phi_n^{(n-1)}(x) dx \\ &= -\phi_m^{(m+1)}(x)\phi_n^{(n-2)}(x)\Big|_{x=-1}^{x=1} + (-1)^2 \int_{-1}^1 \phi_m^{(m+2)}(x)\phi_n^{(n-2)}(x) dx \\ &= \dots \\ &= (-1)^m \int_{-1}^1 \phi_m^{(2m)}(x)\phi_n^{(n-m)}(x) dx \end{aligned}$$

$$= (-1)^m (2m)! \int_{-1}^1 \phi_n^{(n-m)}(x) dx.$$

This is 0 if $m < n$. For $m = n \geq 1$, we have by the change of variable $x = \cos \theta$ that

$$\begin{aligned} \int_{-1}^1 [\phi_n^{(n)}(x)]^2 dx &= (-1)^n (2n)! \int_{-1}^1 (x^2 - 1)^n dx \\ &= 2(2n)! \int_0^1 (1 - x^2)^n dx \\ &= 2(2n)! \int_0^{\pi/2} (\sin \theta)^{2n+1} d\theta \\ &= \frac{2(2n)!(2n)!!}{(2n+1)!!}. \end{aligned}$$

This leads to

$$\begin{aligned} \int_{-1}^1 [P_n(x)]^2 dx &= \frac{1}{2^{2n}(n!)^2} \int_{-1}^1 [\phi_n^{(n)}(x)]^2 dx \\ &= \frac{2(2n)!(2n)!!}{2^{2n}(n!)^2(2n+1)!!} \\ &= \frac{2}{2n+1}. \end{aligned}$$

(2) Let $\tilde{P}_n \in \tilde{\mathcal{P}}_n$ ($n = 0, \dots$) be the unique orthogonal polynomials each with leading coefficient 1 in $L^2(-1, 1)$. By Theorem 1.25, we have

$$\tilde{P}_n(x) = (x - a_n)\tilde{P}_{n-1}(x) - b_n\tilde{P}_{n-2}(x), \quad n = 2, \dots, \quad (1.54)$$

where

$$\begin{aligned} a_n &= \frac{\int_{-1}^1 x[\tilde{P}_{n-1}(x)]^2 dx}{\int_{-1}^1 [\tilde{P}_{n-1}(x)]^2 dx}, \quad n = 1, \dots, \\ b_n &= \frac{\int_{-1}^1 [\tilde{P}_{n-1}(x)]^2 dx}{\int_{-1}^1 [\tilde{P}_{n-2}(x)]^2 dx} \quad n = 2, \dots \end{aligned}$$

By Corollary 1.27, \tilde{P}_n and P_n differ only by a constant. Comparing their leading coefficients, we have

$$\tilde{P}_n = \frac{2^n (n!)^2}{(2n)!} P_n, \quad n = 0, \dots \quad (1.55)$$

In particular, $\tilde{P}_n \in \tilde{\mathcal{P}}_n$ is even (odd) as P_n if n is even (odd). Hence, the function $x[\tilde{P}_{n-1}]^2$ is an odd function, and its integral over $(-1, 1)$ vanishes. Therefore, all $a_n = 0$. By (1.55)

and (1.51),

$$b_n = \frac{(n-1)^2 \int_{-1}^1 [P_{n-1}(x)]^2 dx}{(2n-3)^2 \int_{-1}^1 [P_{n-2}(x)]^2 dx} = \frac{(n-1)^2}{(2n-1)(2n-3)}, \quad n = 2, \dots$$

Consequently, we have by (1.54) with n replaced by $n+1$ that

$$\tilde{P}_{n+1}(x) = x\tilde{P}_n(x) - \frac{n^2}{4n^2-1}\tilde{P}_{n-1}(x), \quad n = 2, \dots$$

This, together with (1.55), leads to (1.52).

(3) This follows from Theorem 1.28.

(4) This follows from Theorem 1.26 and (1.51).

(5) The fact that $P_n(1) = 1$ follows from the three-term recurrence (1.52) with argument of induction. Since P_n is an even (odd) function if n is even (odd), $P_n(-1) = (-1)^n P_n(1) = (-1)^n$.

(6) For $n = 1$, (1.53) is clearly true. We thus assume that $n \geq 2$. Let $q \in \mathcal{P}_{n-1}$. By integration by parts and the fact that P_n is orthogonal to all polynomials in \mathcal{P}_{n-1} , we have

$$\begin{aligned} \int_{-1}^1 [(1-x^2)P_n'(x)]' q(x) dx &= - \int_{-1}^1 (1-x^2)P_n'(x)q'(x) dx \\ &= \int_{-1}^1 [(1-x^2)q'(x)]' P_n(x) dx \\ &= 0. \end{aligned}$$

Therefore, $((1-x^2)P_n'(x))'$ is orthogonal to all polynomials in \mathcal{P}_{n-1} . It thus follows from Corollary 1.27 that

$$\alpha_n P_n(x) = ((1-x^2)P_n'(x))' = (1-x^2)P_n''(x) - 2xP_n'(x).$$

Let c_n be the leading coefficient of P_n . We have

$$\begin{aligned} ((1-x^2)P_n'(x))' &= c_n [(1-x^2)P_n''(x) - 2xP_n'(x)] \\ &= c_n \{(1-x^2)[n(n-1)x^{n-2} + \dots] - 2x(nx^{n-1} + \dots)\} \\ &= c_n n(n+1)(x^n + \dots). \end{aligned}$$

These two equations imply (1.53) for $n \geq 2$. □

Exercises

1. Let $B_n f \in \mathcal{P}_n$ ($n = 0, \dots$) be the Bernstein polynomials of $f \in C[0, 1]$.

(a) Let $f_0(x) = 1$, $f_1(x) = x$, and $f_2(x) = x^2$. Show that

$$B_n f_0(x) = 1, \quad B_n f_1(x) = x, \quad B_n f_2(x) = \frac{n-1}{n}x^2 + \frac{1}{n}x, \quad \forall x \in [0, 1].$$

(b) In general, is $B_n f \in \mathcal{P}_n$ the best uniform approximation of $f \in C[0, 1]$ in \mathcal{P}_n on $[0, 1]$?

2. Let $B_n f \in \mathcal{P}_n$ ($n = 0, \dots$) be the Bernstein polynomials of $f \in C[0, 1]$. Prove that $\|(B_n f)' - f'\|_{C[a,b]} \rightarrow 0$ for any $f \in C^1[0, 1]$.
3. Let $0 < a < b < 1$ and $f \in C[a, b]$. Show that there exist a sequence of polynomials with integer coefficients that converge to f in the $C[a, b]$ -norm.
4. Denote $\|f\| = \|f\|_{C[a,b]}$. Show that for any $f, g \in C[a, b]$

$$\|f + g\| \leq \|f\| + \|g\| \quad \text{and} \quad \left| \|f\| - \|g\| \right| \leq \|f - g\|.$$

5. Prove Theorem 1.3.

6. Let $n \geq 0$ be an integer and a, b two real numbers with $a < b$. Define

$$F(c) = \max_{a \leq x \leq b} \left| \sum_{k=0}^n c_k x^k \right| \quad \forall c = (c_0, \dots, c_n) \in \mathbb{R}^{n+1}.$$

Show that $F : \mathbb{R}^{n+1} \rightarrow \mathbb{R}$ is a continuous function.

7. Let $k \geq 1$ be an integer and $f \in C^k[a, b]$. Show that, for any $\epsilon > 0$, there exists $p \in \mathcal{P}$ such that

$$\|f - p\|_{C[a,b]} < \epsilon, \quad \|f' - p'\|_{C[a,b]} < \epsilon, \quad \dots, \quad \|f^{(k)} - p^{(k)}\|_{C[a,b]} < \epsilon.$$

8. Let $f \in C[a, b]$ but $f \notin \mathcal{P}$. Show that there exists no polynomial $p \in \mathcal{P}$ such that

$$\|f - p\|_{C[a,b]} \leq \|f - q\|_{C[a,b]} \quad \forall q \in \mathcal{P}.$$

9. Let x_0, \dots, x_n be $n + 1$ distinct points in $[a, b]$. Let $f \in C[a, b]$. Does there exist a unique $p \in \mathcal{P}_m$ such that

$$\max_{0 \leq j \leq n} |f(x_j) - p(x_j)| \leq \max_{0 \leq j \leq n} |f(x_j) - q(x_j)| \quad \forall q \in \mathcal{P}_m?$$

Discuss the cases $0 \leq m < n$, $m = n$, and $m > n$.

10. Let f_1, \dots, f_n be n linearly independent functions in $C[a, b]$. Let $f \in C[a, b]$. Show that there exists $p \in \text{Span}\{f_1, \dots, f_n\}$ such that

$$\|f - p\|_{C[a,b]} \leq \|f - q\|_{C[a,b]} \quad \forall q \in \text{Span}\{f_1, \dots, f_n\}.$$

Discuss the uniqueness of such an approximation.

11. Let $f \in C[a, b]$ and $q_n \in \mathcal{P}_n$ for some integer $n \geq 0$. Suppose $p_n \in \mathcal{P}_n$ is the best uniform approximation of f in \mathcal{P}_n . Prove that $p_n + q_n$ is the best uniform approximation of $f + q_n$ in \mathcal{P}_n .
12. Let $c > 0$. Let $f \in C[-c, c]$ be an even (odd) function. Show that the best uniform approximation of f in \mathcal{P}_n for an integer $n \geq 0$ is also an even (odd) function.
13. Show that $p_1(x) = x - 1/8$ is the best uniform approximation of $f(x) = x^2$ in \mathcal{P}_1 on $[0, 1]$.
14. Let $f(x) = x^4$ ($0 \leq x \leq 1$). Find the best uniform approximation of f in \mathcal{P}_1 on $[0, 1]$.
15. Find the best uniform approximation of x^{n+2} in \mathcal{P}_n with respect to the $C[-1, 1]$ -norm.
16. Let $f \in C[a, b]$ but $f \notin \mathcal{P}_n$ for some integer $n \geq 0$. Let $p \in \mathcal{P}_n$ be the best uniform approximation of f in \mathcal{P}_n on $[a, b]$. Can there exist a sequence of strictly increasing points $\{x_k\}_{k=1}^\infty$ in $[a, b]$ such that

$$|f(x_k) - p(x_k)| = \|f - p\|_{C[a,b]} \quad \forall k \geq 1?$$

or such that

$$f(x_k) - p(x_k) = (-1)^k \|f - p\|_{C[a,b]} \quad \forall k \geq 1?$$

17. Let $n \geq 1$ be an integer. Denote by $\hat{\mathcal{T}}_n$ the set of all functions

$$\hat{T}(x) = a_0 + \sum_{k=1}^n (a_k \cos^k x + b_k \sin^k x)$$

with a_0, a_1, \dots, a_n and b_1, \dots, b_n all real numbers. Show that $\mathcal{T}_n = \hat{\mathcal{T}}_n$.

18. Show that any nonzero trigonometric polynomial of degree less than or equal to n can have at most $2n$ zeros in $[0, 2\pi)$.
19. Prove the Second Weierstrass Approximation Theorem by the First Weierstrass Approximation Theorem.
20. Show that the best uniform approximation of an even (odd) function $f \in C_{2\pi}$ in \mathcal{T}_n is also an even (odd) function.
21. Given any function g on $[a, b]$, define

$$g^*(\theta) = g\left(\frac{(b-a)\cos\theta + (a+b)}{2}\right) \quad \forall \theta \in (-\infty, \infty).$$

Let $f \in C[a, b]$ and $n \geq 0$ be an integer. Let $p \in \mathcal{P}_n$ and $T \in \mathcal{T}_n$ satisfy, respectively,

$$\|f - p\|_{C[a,b]} = E_n(f) := \min_{q \in \mathcal{P}_n} \|f - q\|_{C[a,b]}$$

and

$$\|f^* - T\|_{C_{2\pi}} = E_n^*(f^*) := \min_{S \in \mathcal{T}_n} \|f^* - S\|_{C_{2\pi}}.$$

Show that $E_n(f) = E_n^*(f^*)$ and that $T = p^*$.

22. Let $f \in C[a, b]$. Let ω_f be the modulus of continuity of f over $[a, b]$. Show for any integer $k \geq 1$ that

$$\omega_f(k\delta) \leq k\omega_f(\delta)$$

and for any $\lambda > 0$ that

$$\omega_f(\lambda\delta) \leq (\lambda + 1)\omega_f(\delta).$$

23. Let $f \in C(-\infty, \infty)$. Define for any $\delta > 0$

$$R_f(\delta) = \frac{1}{\delta} \int_{x-\frac{\delta}{2}}^{x+\frac{\delta}{2}} f(t) dt.$$

Show that

$$|f(x) - R_f(\delta)| \leq \omega_f(\delta) \quad \forall x \in (-\infty, \infty), \delta > 0,$$

where $\omega_f(\delta)$ is the modulus of continuity of f .

24. Show that there exists a constant $K = K(a, b) > 0$ independent of f such that

$$E_n(f) \leq \frac{K}{n} E_{n-1}(f') \quad \forall f \in C^1[a, b], \forall n \geq 1,$$

where for any integer $k \geq 0$

$$E_k(f) = \min_{q \in \mathcal{P}_k} \|f - q\|_{C[a, b]}.$$

Let $f \in C^p[a, b]$ for some integer $p \geq 1$. Show that there exists a constant $K = K(a, b, p) > 0$ independent of f such that

$$E_n(f) \leq \frac{K \|f^{(p)}\|_{C[a, b]}}{n^p} \quad \forall n \geq p + 1,$$

25. Show that

$$\frac{2}{\pi} < \frac{\sin t}{t} < 1 \quad \forall t \in \left(0, \frac{\pi}{2}\right).$$

26. Let $f \in C_{2\pi}$. Let

$$a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos kx dx \quad \text{and} \quad b_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin kx dx$$

be the Fourier coefficients of f . Let

$$S_n f(x) = \frac{a_0}{2} + \sum_{k=1}^n (a_k \cos kx + b_k \sin kx)$$

be the partial sum of the Fourier series of f for any integer $n \geq 1$.

(a) Show that

$$S_n f(x) = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t+x) \frac{\sin\left(n + \frac{1}{2}\right)t}{2 \sin \frac{t}{2}} dt$$

and that

$$\|S_n f\|_{C_{2\pi}} \leq \lambda_n \|f\|_{C_{2\pi}},$$

where

$$\lambda_n = \frac{1}{\pi} \int_{-\pi}^{\pi} \left| \frac{\sin\left(n + \frac{1}{2}\right)t}{2 \sin \frac{t}{2}} \right| dt.$$

(b) Show that

$$\|f - S_n f\|_{C_{2\pi}} \leq (1 + \lambda_n) E_n^*(f),$$

where

$$E_n^*(f) = \min_{T \in \mathcal{T}_n} \|f - T\|_{C_{2\pi}}.$$

Show also that

$$\frac{4 \log n}{\pi^2} < \lambda_n < 2 + \log n.$$

(c) Show that

$$\|f - S_n f\|_{C_{2\pi}} \leq (3 + \log n) E_n^*(f).$$

27. Prove that

$$T_m(T_n(x)) = T_{mn}(x) \quad \forall m, n \geq 0.$$

28. Prove that

$$\int_{-1}^1 [T_n(x)]^2 = 1 - \frac{1}{4n^2 - 1} \quad \forall n \geq 0.$$

29. Calculate $T'_n(\pm 1)$ for any integer $n \geq 0$.

30. (Chebyshev) Let $n \geq 0$ be an integer and T_n the n th Chebyshev polynomial of first kind. Let $P \in \mathcal{P}_n$ satisfy that $|P(x)| \leq 1$ for all $x \in [-1, 1]$. Show that

$$|P(y)| \leq |T_n(y)| \quad \forall y \notin [-1, 1].$$

31. Prove the following properties of the Chebyshev Polynomials of second kind

$$U_n(x) = \frac{\sin(n+1)\theta}{\sin \theta}, \quad n = 0, \dots,$$

where $x = \cos \theta$ ($\theta \in [0, \pi]$):

(a) *Recursion formula.*

$$\begin{aligned} U_0(x) &= 1, & U_1(x) &= 2x, \\ U_{n+1}(x) &= 2xU_n(x) - U_{n-1}(x), & n &= 1, \dots; \end{aligned}$$

(b) *Orthogonality.*

$$\int_{-1}^1 \sqrt{1-x^2} U_m(x) U_n(x) dx = \begin{cases} 0 & \text{if } m \neq n, \\ \pi/2 & \text{if } m = n; \end{cases}$$

(c) *Differential equations.*

$$(1-x^2) U_n''(x) - 3x U_n'(x) + n(n+2) U_n(x) = 0 \quad n = 0, \dots;$$

(d) *Relations with the Chebyshev polynomials of first kind.*

$$\begin{aligned} n U_{n-1}(x) &= T_n'(x), & n &= 1, \dots, \\ U_n(x) &= x U_{n-1}(x) + T_n(x), & n &= 1, \dots; \end{aligned}$$

(e) For each $n \geq 0$, U_n is a polynomial of degree n with leading coefficient 2^n .
Moreover, if n is even (odd), then U_n is an even (odd) polynomial.

32. Define $\chi(x) = -1$ if $-1 \leq x < 0$ and $\chi(x) = 1$ if $0 \leq x \leq 1$.

(a) Show that

$$\inf_{f \in C[-1,1]} \sup_{-1 \leq x \leq 1} |f(x) - \chi(x)| = 1,$$

and that there exist infinitely many $f \in C[-1, 1]$ such that

$$\sup_{-1 \leq x \leq 1} |f(x) - \chi(x)| = 1.$$

(b) Show that

$$\inf_{f \in C[-1,1]} \int_{-1}^1 |f(x) - \chi(x)|^2 dx = 0,$$

and that there exists no $f \in C[-1, 1]$ such that

$$\int_{-1}^1 |f(x) - \chi(x)|^2 dx = 0.$$

33. Show that

$$\int_0^1 |f(x)| dx \leq \sqrt{\int_0^1 |f(x)|^2 dx} \quad \forall f \in C[0, 1].$$

34. Let \mathfrak{S} be an inner product space. Let f_1, \dots, f_n be n linearly independent vectors in \mathfrak{S} and $\mathfrak{S}_n = \text{Span}\{f_1, \dots, f_n\}$. Let $f \in \mathfrak{S}$. Prove the following:

(a) There exists a unique $p \in \mathfrak{S}_n$ such that

$$\|f - p\| = \min_{q \in \mathfrak{S}_n} \|f - q\|.$$

This p is called the *least-squares approximation* of f in \mathfrak{S}_n ;

(b) The least-squares approximation $p \in \mathcal{S}_n$ of f is characterized by

$$\langle f - p, q \rangle = 0 \quad \forall q \in \mathcal{S}_n;$$

(c) The error of the least-squares approximation is given by

$$\|f - p\|^2 = \|f\|^2 - \|p\|^2.$$

35. Let $\{f_n\}_{n=1}^{\infty}$ be an orthonormal system of an inner product space \mathcal{S} . Prove the Bessel inequality

$$\sum_{n=1}^{\infty} \langle f, f_n \rangle^2 \leq \|f\|^2 \quad \forall f \in \mathcal{S}.$$

36. Find the least-squares approximation of $f(x) = x^3$ in \mathcal{P}_1 over $[-1, 1]$.

37. Find the least-squares approximation of $f(x) = x^4$ in \mathcal{P}_1 over $[0, 1]$.

38. Let $p(x) = \sum_{k=0}^n a_k x^k \in \mathcal{P}_n$ be the least squares approximation of a given $f \in C[0, 1]$ over $[0, 1]$. Find the coefficient matrix of the normal equation.

39. Let $f \in C[a, b]$ and define

$$\mu_n(f) = \int_a^b x^n f(x) dx, \quad n = 0, \dots$$

Show that $f(x) = 0$ for all $x \in [a, b]$ if and only $\mu_n(f) = 0$ for all $n = 0, \dots$

40. Let

$$P_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} [(x^2 - 1)^n], \quad n = 0, \dots$$

be the Legendre polynomials.

(a) Let $n \geq 1$. Prove directly by Rolle's Theorem that P_n has n simple roots in $(-1, 1)$.

(b) Let $r \geq 1$ be an integer. Show that the sequence of corresponding derivatives $\{P_n^{(r)}\}_{n=1}^{\infty}$ is orthogonal with respect to the weight function $(1 - x^2)^r$, i.e.,

$$\int_{-1}^1 P_m^{(r)}(x) P_n^{(r)}(x) (1 - x^2)^r dx = 0 \quad \text{if } m \neq n.$$

41. Let $n \geq 0$ be an integer and P_n be defined as in the previous problem.

(a) Show that P'_{n+1} has n distinct roots ξ_1, \dots, ξ_n in $(-1, 1)$.

(b) Show that

$$\frac{d}{dx} \left[x^{n+1} - \sum_{k=0}^n \left(\int_{-1}^1 t^{n+1} P_k(t) dt \right) P_k(x) \right]$$

vanishes at these points ξ_1, \dots, ξ_n .

42. Let $\{Q_n\}_{n=0}^{\infty}$ be an orthonormal system of polynomials in $L^2_{\rho}(a, b)$. Prove for any $n \geq 0$ the identity

$$\sum_{k=0}^n [Q_k(x)]^2 = \frac{\alpha_n}{\alpha_{n+1}} [Q'_{n+1}(x)Q_n(x) - Q'_n(x)Q_{n+1}(x)],$$

where α_k is the leading coefficient of Q_k ($k = 0, \dots$).

43. Let $\{Q_n\}_{n=0}^{\infty}$ be an orthogonal system of polynomials in $L^2_{\rho}(a, b)$. Let $n \geq 1$. Prove that the zeros of Q_n and that of Q_{n+1} alternate.

Chapter 2

Polynomial Interpolation

2.1 Lagrange Interpolation

Let $n \geq 0$ be an integer, x_0, \dots, x_n distinct points in a finite interval $[a, b]$, and $y_0, \dots, y_n \in \mathbb{R}$.

Definition 2.1 (Larange interpolation). A Lagrange interpolation polynomial, or Lagrange interpolant, that interpolates y_0, \dots, y_n at x_0, \dots, x_n is a polynomial $p_n \in \mathcal{P}_n$ such that

$$p_n(x_i) = y_i, \quad i = 0, \dots, n. \quad (2.1)$$

The points x_0, \dots, x_n are called the interpolation points. In the case $y_i = f(x_i)$ ($i = 0, \dots, n$) for some function $f : [a, b] \rightarrow \mathbb{R}$, p_n is also called a Lagrange interpolation polynomial, or Lagrange interpolant, of f at x_0, \dots, x_n .

Theorem 2.2 (Existence and uniqueness of Lagrange interpolation). There exists a unique Lagrange interpolation polynomial that interpolates y_0, \dots, y_n at x_0, \dots, x_n .

Proof. If $n = 0$, then the unique $p_0 \in \mathcal{P}_0$ is given by $p_0(x) = y_0$. Let $n \geq 1$. Consider a general polynomial in \mathcal{P}_n

$$p_n(x) = a_0 + a_1x + \dots + a_nx^n,$$

where $a_i \in \mathbb{R}$ ($i = 0, \dots, n$). Eq. (2.1) is equivalent to the system of linear equations of the unknowns a_0, \dots, a_n

$$a_0 + a_1x_i + \dots + a_nx_i^n = f(x_i), \quad i = 0, \dots, n. \quad (2.2)$$

The determinant of the coefficient matrix of this linear system is the Vandermonde determinant

$$V(x_0, x_1, \dots, x_n) = \begin{vmatrix} 1 & x_0 & x_0^2 & \dots & x_0^n \\ 1 & x_1 & x_1^2 & \dots & x_1^n \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_n & x_n^2 & \dots & x_n^n \end{vmatrix} = \prod_{0 \leq i < j \leq n} (x_j - x_i).$$

This is nonzero, since x_0, \dots, x_n are distinct. Therefore, (2.2) has a unique solution. This implies the desired existence and uniqueness. \square

For $n \geq 1$, we define

$$\begin{aligned} l_i(x) &= \frac{(x - x_0) \cdots (x - x_{i-1})(x - x_{i+1}) \cdots (x - x_n)}{(x_i - x_0) \cdots (x_i - x_{i-1})(x_i - x_{i+1}) \cdots (x_i - x_n)} \\ &= \prod_{\substack{j=0 \\ j \neq i}}^n \frac{x - x_j}{x_i - x_j}, \quad i = 0, \dots, n. \end{aligned} \quad (2.3)$$

Clearly, each $l_i \in \mathcal{P}_n$ has the degree exactly n . Moreover,

$$l_i(x_j) = \delta_{ij}, \quad i, j = 0, \dots, n. \quad (2.4)$$

If $n = 0$, we define $l_0(x) = 1$.

Theorem 2.3 (Lagrange's formula of Lagrange interpolation). *The unique Lagrange interpolation polynomial $p_n \in \mathcal{P}_n$ that interpolates y_0, \dots, y_n at x_0, \dots, x_n is given by*

$$p_n(x) = y_0 l_0(x) + y_1 l_1(x) + \cdots + y_n l_n(x) = \sum_{i=0}^n y_i l_i(x). \quad (2.5)$$

Proof. Let p_n be given by (2.5). Clearly, $p_n \in \mathcal{P}_n$. Moreover, by (2.4),

$$p_n(x_j) = \sum_{i=0}^n y_i l_i(x_j) = \sum_{i=0}^n y_i \delta_{ij} = y_j, \quad j = 0, \dots, n.$$

Thus, p_n is the Lagrange interpolation polynomial in \mathcal{P}_n . \square

The formula (2.1) is called *Lagrange's formula* of the Lagrange interpolation.

Example. Find the Lagrange interpolation polynomial $p_2 \in \mathcal{P}_2$ that interpolates $y_0 = -1, y_1 = 3, y_2 = 2$ at $x_0 = 0, x_1 = 1, x_2 = 2$.

We first calculate the polynomials l_0, l_1 , and l_2 associated with the points x_0, x_1 , and x_2 .

$$\begin{aligned} l_0(x) &= \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)} = \frac{(x - 1)(x - 2)}{(0 - 1)(0 - 2)} = \frac{1}{2}x^2 - \frac{3}{2}x + 1, \\ l_1(x) &= \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)} = \frac{(x - 0)(x - 2)}{(1 - 0)(1 - 2)} = -x^2 + 2x, \\ l_2(x) &= \frac{(x - x_0)(x - x_1)}{(x_2 - x_0)(x_2 - x_1)} = \frac{(x - 0)(x - 1)}{(2 - 0)(2 - 1)} = \frac{1}{2}x^2 - \frac{1}{2}. \end{aligned}$$

We have now by Lagrange's formula (2.1) that

$$\begin{aligned} p_2(x) &= y_0 l_0(x) + y_1 l_1(x) + y_2 l_2(x) \\ &= (-1) \left(\frac{1}{2}x^2 - \frac{3}{2}x + 1 \right) + 3(-x^2 + 2x) + 2 \left(\frac{1}{2}x^2 - \frac{1}{2} \right) \\ &= -\frac{5}{2}x^2 + \frac{13}{2}x - 1. \end{aligned}$$

We can check that

$$\begin{aligned} p_2(0) &= -\frac{5}{2} \cdot 0^2 + \frac{13}{2} \cdot 0 - 1 = -1, \\ p_2(1) &= -\frac{5}{2} \cdot 1^2 + \frac{13}{2} \cdot 1 - 1 = 3, \\ p_2(2) &= -\frac{5}{2} \cdot 2^2 + \frac{13}{2} \cdot 2 - 1 = 2. \end{aligned}$$

The polynomials l_0, \dots, l_n defined in (2.3) are called the *Lagrange basis polynomials* associated with the $n+1$ distinct points x_0, \dots, x_n . The word "basis" is justified in the first part of the following proposition:

Proposition 2.4. (1) *The polynomials l_0, \dots, l_n defined in (2.3) form a basis of \mathcal{P}_n .*
(2) *For any $p_n \in \mathcal{P}_n$,*

$$\sum_{i=0}^n p_n(x_i) l_i(x) = p_n(x) \quad \forall x \in \mathbb{R}, \quad (2.6)$$

$$\sum_{i=0}^n p_n(x - x_i) l_i(x) = p_n(0) \quad \forall x \in \mathbb{R}. \quad (2.7)$$

Proof. (1) Let $c_0, \dots, c_n \in \mathbb{R}$ satisfy $\sum_{i=0}^n c_i l_i = 0$ in \mathcal{P}_n , i.e., $\sum_{i=0}^n c_i l_i(x) = 0$ for all $x \in \mathbb{R}$. Setting $x = x_j$ for an arbitrary j with $0 \leq j \leq n$, we obtain by (2.4) that $c_j = 0$. Thus, l_0, \dots, l_n are linearly independent in \mathcal{P}_n , and form a basis of \mathcal{P}_n , since $\dim \mathcal{P}_n = n+1$.

(2) Let $q_n(x)$ denote the left-hand side of the identity in (2.6). Clearly, $q_n \in \mathcal{P}_n$. Moreover, it follows from (2.4) that $q_n(x_j) = p_n(x_j)$ for all $j = 0, \dots, n$. Thus, the polynomial $p_n - q_n \in \mathcal{P}_n$ vanishes at $n+1$ distinct points. Since any nonzero polynomial in \mathcal{P}_n can have at most n zeros, $q_n(x) - p_n(x)$ must be identically zero. This proves (2.6).

To prove (2.7), we fix an arbitrary $x \in \mathbb{R}$. Notice that $p_n(x - \cdot) \in \mathcal{P}_n$. Thus, by (2.6),

$$\sum_{i=0}^n p_n(x - x_i) l_i(t) = p_n(x - t) \quad \forall t \in \mathbb{R}.$$

Setting $t = x$, we obtain (2.7). □

For any $f \in C[a, b]$, we denote by $L_n f \in \mathcal{P}_n$ the Lagrange interpolation polynomial that interpolates f at x_0, \dots, x_n . We call $L_n : C[a, b] \rightarrow \mathcal{P}_n$ the *Lagrange interpolation operator*, or simply *Lagrange interpolator*, associated with the $n + 1$ distinct points $x_0, \dots, x_n \in [a, b]$.

Proposition 2.5. (1) *Each $L_n : C[a, b] \rightarrow \mathcal{P}_n$ is a linear operator.*
(2) *$L_n f = f$ for any $f \in \mathcal{P}_n$.*

Proof. (1) This follows from Lagrange's formula (2.5).

(2) This follows from (2.5) and (2.6). □

For each $k \geq 1$, we denote by $C^k[a, b]$ the set of functions $f : [a, b] \rightarrow \mathbb{R}$ that have all the continuous derivatives $f^{(j)}$ on $[a, b]$ for $1 \leq j \leq k$. The derivatives at the end-points a and b are the one-sided derivatives, and the continuity at a and b is also one-sided. For convenience, we denote $C^0[a, b] = C[a, b]$.

Theorem 2.6 (The remainder of Lagrange interpolation). *Let x_0, \dots, x_n be $n + 1$ distinct points in $[a, b]$. Let $f \in C^{n+1}[a, b]$ and $L_n f \in \mathcal{P}_n$ be the Lagrange interpolant of f at x_0, \dots, x_n . Then for any $x \in [a, b]$ there exists $\xi(x) \in [a, b]$ such that*

$$f(x) - (L_n f)(x) = \frac{f^{(n+1)}(\xi(x))}{(n+1)!} (x - x_0) \cdots (x - x_n) \quad \forall x \in [a, b]. \quad (2.8)$$

Proof. Let $x \in [a, b]$. If $x = x_i$ for some i with $0 \leq i \leq n$, then (2.8) holds true for any $\xi(x) \in [a, b]$. Assume that $x \neq x_i$ ($0 \leq i \leq n$). Let $\omega(t) = \prod_{i=0}^n (t - x_i)$ and define

$$\phi(t) = f(t) - (L_n f)(t) - \lambda \omega(t),$$

where $\lambda \in \mathbb{R}$ is so chosen that $\phi(x) = 0$, i.e.,

$$\lambda = \frac{f(x) - (L_n f)(x)}{\omega(x)}. \quad (2.9)$$

Clearly $\phi \in C^{n+1}[a, b]$. Moreover, $\phi = 0$ at the $n + 2$ distinct points x, x_0, \dots, x_n in $[a, b]$. Thus, by Rolle's Theorem,

$$\phi' = 0 \text{ at } n + 1 \text{ distinct points in } [a, b],$$

$$\phi'' = 0 \text{ at } n \text{ distinct points in } [a, b],$$

...

$$\phi^{(n)} = 0 \text{ at } 2 \text{ distinct points in } [a, b].$$

Finally, there exists $\xi(x) \in [a, b]$ such that $\phi^{(n+1)}(\xi(x)) = 0$. By the definition of ϕ , we have

$$\phi^{(n+1)}(t) = f^{(n+1)}(t) - \lambda(n+1)!$$

Hence,

$$\phi^{(n+1)}(\xi(x)) = f^{(n+1)}(\xi(x)) - \lambda(n+1)! = 0$$

This, together with (2.9), implies (2.8). □

Consider now the special case $[a, b] = [-1, 1]$. Let $f \in C^{n+1}[-1, 1]$. By the above theorem, we have

$$|f(x) - (L_n f)(x)| \leq \frac{1}{(n+1)!} \left(\max_{a \leq x \leq b} |f^{(n+1)}(x)| \right) |\omega(x)| \quad \forall x \in [-1, 1], \quad (2.10)$$

where $L_n : C[-1, 1] \rightarrow \mathcal{P}_n$ is the Lagrange interpolator associated with a given set of $n+1$ distinct points x_0, \dots, x_n in $[-1, 1]$ and $\omega(x) = \prod_{k=0}^n (x - x_k)$.

In order to minimize the error $f - L_n f$ for all $f \in C^{n+1}[-1, 1]$ with respect to the $C[-1, 1]$ -norm, we choose $x_0, \dots, x_n \in [-1, 1]$ to minimize the $C[-1, 1]$ -norm of ω . Since $\omega \in \tilde{\mathcal{P}}_n$, it follows from Theorem 1.10 on properties of Chebyshev polynomials that the optimal choice of ω is the rescaled Chebyshev polynomial:

$$\omega(x) = \tilde{T}_{n+1}(x) = 2^{-n} T_{n+1}(x) \quad \forall x \in [-1, 1]. \quad (2.11)$$

In particular, the unique set of optimal interpolation points are the roots of Chebyshev polynomials T_{n+1} :

$$x_k = \cos \frac{(2k+1)\pi}{2(n+1)}, \quad k = 0, \dots, n.$$

Moreover,

$$\|\omega\|_{C[-1,1]} = \|\tilde{T}_{n+1}\|_{C[-1,1]} = \|2^{-n} T_{n+1}\|_{C[-1,1]} = 2^{-n}. \quad (2.12)$$

Therefore, with such a choice of interpolation points, we have the error estimate

$$\|f - L_n f\|_{C[-1,1]} \leq \frac{1}{2^n (n+1)!} \|f^{(n+1)}\| \quad \forall f \in C^{n+1}[-1, 1]. \quad (2.13)$$

If we set $f = \tilde{T}_{n+1}$, then the unique Lagrange interpolation polynomial in \mathcal{P}_n of f at all the roots of T_{n+1} is the zero polynomial. By (2.12), the equality in (2.13) holds true. Hence, the error estimate (2.13) is optimal.

Now consider the interpolation error in the $L_\rho^2(a, b)$ -norm for some weight function ρ on $[a, b]$. Let $f \in C^{n+1}[a, b]$. By (2.10), we have

$$\|f - L_n f\|_{L_\rho^2(a,b)} \leq \frac{1}{(n+1)!} \|f^{(n+1)}\|_{C[a,b]} \|\omega\|_{L_\rho^2(a,b)}, \quad (2.14)$$

where again $L_n : C[a, b] \rightarrow \mathcal{P}_n$ is the Lagrange interpolator associated with a given set of $n+1$ distinct points $x_0, \dots, x_n \in [a, b]$ and $\omega(x) = \prod_{k=0}^n (x - x_k)$. By Theorem 1.26, the unique set of optimal interpolation points x_0, \dots, x_n are the $n+1$ distinct roots of the $(n+1)$ st orthogonal polynomial $Q_{n+1} \in \tilde{\mathcal{P}}_{n+1}$ in $L_\rho^2(a, b)$ i.e., $\omega = Q_{n+1}$. Thus, it follows from (2.14) that

$$\|f - L_n f\|_{L_\rho^2(a,b)} \leq \frac{1}{(n+1)!} \|f^{(n+1)}\|_{C[a,b]} \|Q_{n+1}\|_{L_\rho^2(a,b)} \quad \forall f \in C^{n+1}[a, b].$$

Taking $f = Q_{n+1}$, we have $L_n f = 0$. Hence, this estimate is optimal.

In the special case $[a, b] = [-1, 1]$ and $\rho(x) = 1/\sqrt{1-x^2}$, we see from Theorem 1.10 that the unique set of optimal interpolation points in $[-1, 1]$ are the zeros of Chebyshev polynomial T_{n+1} and that $\omega = 2^{-n}T_{n+1}$. By (1.27), we have

$$\|\omega\|_{L^2_\rho(-1,1)} = \|2^{n+1}T_{n+1}\|_{L^2_\rho(-1,1)} = \frac{\sqrt{\pi}}{2^{n+1/2}}.$$

Therefore, we have the optimal estimate

$$\|f - L_n f\|_{L^2_\rho(-1,1)} \leq \frac{\sqrt{\pi}}{2^{n+1/2}(n+1)!} \|f^{(n+1)}\|_{C[-1,1]} \quad \forall f \in C^{n+1}[-1, 1].$$

In the special case $[a, b] = [-1, 1]$ and $\rho(x) = 1$, the unique set of optimal interpolation points in $[-1, 1]$ are the zeros of Legendre polynomial P_{n+1} and that $\omega = [2^n(n!)^2/(2n)!]P_{n+1}$. By (1.51), we have

$$\|\omega\|_{L^2(-1,1)} = \frac{2^n(n!)^2}{(2n)!} \sqrt{\frac{2}{2n+1}}.$$

Therefore, we have the optimal estimate

$$\|f - L_n f\|_{L^2_\rho(-1,1)} \leq \frac{2^n n!}{(n+1)(2n)!} \sqrt{\frac{2}{2n+1}} \|f^{(n+1)}\|_{C[-1,1]} \quad \forall f \in C^{n+1}[-1, 1].$$

2.2 Newton's Formula and Divided Differences

Suppose $p_k \in \mathcal{P}_k$ is the Lagrange interpolation polynomial that interpolates f_0, \dots, f_k at x_0, \dots, x_k . Consider adding one more interpolation point $x_{k+1} \in \mathbb{R}$ that is different from all x_0, \dots, x_k , and adding one more value $f_{k+1} \in \mathbb{R}$. Let $p_{k+1} \in \mathcal{P}_{k+1}$ be the Lagrange interpolation polynomial $p_{k+1} \in \mathcal{P}_{k+1}$ that interpolates f_0, \dots, f_k , and f_{k+1} at x_0, \dots, x_k , and x_{k+1} . Since $p_k(x_i) = p_{k+1}(x_i) = f_i$ for $i = 0, \dots, k$, we see that the polynomial $p_{k+1} - p_k \in \mathcal{P}_{k+1}$ vanishes at x_0, \dots, x_k . Hence, it must have the form

$$p_{k+1}(x) - p_k(x) = d_{k+1}(x - x_0) \cdots (x - x_k)$$

for some $d_{k+1} \in \mathbb{R}$. The condition that $p_{k+1}(x_{k+1}) = f_{k+1}$ determines uniquely that

$$d_{k+1} = \frac{f_{k+1} - p_k(x_{k+1})}{(x_{k+1} - x_0) \cdots (x_{k+1} - x_k)}. \quad (2.15)$$

Therefore, starting from the constant polynomial $p_0(x) = d_0 \in \mathbb{R}$ that interpolates f_0 at x_0 , we have

$$p_0(x) = d_0,$$

$$\begin{aligned}
p_1(x) &= p_0(x) + d_1(x - x_0), \\
p_2(x) &= p_1(x) + d_2(x - x_0)(x - x_1), \\
&\dots \\
p_k(x) &= p_{k-1}(x) + d_k(x - x_0) \cdots (x - x_{k-1}),
\end{aligned}$$

where $d_0, \dots, d_k \in \mathbb{R}$ are constants. Finally,

$$p_k(x) = d_0 + d_1(x - x_0) + \cdots + d_k(x - x_0) \cdots (x - x_{k-1}).$$

Definition 2.7 (Divided differences). *The divided differences of a given set of numbers f_0, \dots, f_n at $n + 1$ distinct points $x_0, \dots, x_n \in \mathbb{R}$ are*

$$\begin{aligned}
f[x_0] &= f_0, \\
f[x_0, \dots, x_k] &= \frac{f[x_1, \dots, x_k] - f[x_0, \dots, x_{k-1}]}{x_k - x_0}, \quad k = 2, \dots, n.
\end{aligned}$$

If $f_i = f(x_i)$ ($i = 0, \dots, n$) for some function f that is defined on a set of real numbers containing all x_0, \dots, x_n , then $f[x_0], \dots, f[x_0, \dots, x_n]$ are called the divided differences of the function f at these points x_0, \dots, x_n .

Theorem 2.8 (Newton's formula of Lagrange interpolation). *Let $x_0, \dots, x_n \in \mathbb{R}$ be $n + 1$ distinct points and $f_0, \dots, f_n \in \mathbb{R}$. Then, for each integer k with $0 \leq k \leq n$,*

$$p_k(x) = f[x_0] + f[x_0, x_1](x - x_0) + \cdots + f[x_0, \dots, x_k](x - x_0) \cdots (x - x_{k-1}) \quad (2.16)$$

is the Lagrange interpolation polynomial that interpolates f_0, \dots, f_k at x_0, \dots, x_k .

The formula (2.16) is called the *Newton's formula* of the Lagrange interpolation.

Example. Use Newton's formula to find the Lagrange interpolation polynomial $p_2 \in \mathcal{P}_2$ that interpolates $f_0 = -1, f_1 = 3, f_2 = 2$ at $x_0 = 0, x_1 = 1, x_2 = 2$.

We first calculate all the needed divided differences.

$$\begin{aligned}
f[x_0] &= f_0 = -1, \\
f[x_1] &= f_1 = 3, \\
f[x_2] &= f_2 = 2, \\
f[x_0, x_1] &= \frac{f[x_1] - f[x_0]}{x_1 - x_0} = \frac{3 - (-1)}{1 - 0} = 4, \\
f[x_1, x_2] &= \frac{f[x_2] - f[x_1]}{x_2 - x_1} = \frac{2 - 3}{2 - 1} = -1, \\
f[x_1, x_1, x_2] &= \frac{f[x_1, x_2] - f[x_0, x_1]}{x_2 - x_0} = \frac{-1 - 4}{2 - 0} = -\frac{5}{2}.
\end{aligned}$$

By Newton's formula (2.16), we have

$$\begin{aligned} p_2(x) &= f[x_0] + f[x_0, x_1](x - x_0) + f[x_0, x_1, x_2](x - x_0)(x - x_1) \\ &= -1 + 4(x - 0) + \left(-\frac{5}{2}\right)(x - 0)(x - 1) \\ &= -\frac{5}{2}x^2 + \frac{13}{2}x - 1. \end{aligned}$$

This is the same polynomial as obtained in the example in Section 2.1.

To prove Theorem 2.8, we first prove the following useful lemma:

Lemma 2.9. *Suppose $p_k, q_k \in \mathcal{P}_k$ are the Lagrange interpolation polynomials that interpolate f_0, \dots, f_k at x_0, \dots, x_k and f_1, \dots, f_{k+1} at x_1, \dots, x_{k+1} , respectively. Then,*

$$r_{k+1}(x) = \frac{(x - x_0)q_k(x) - (x - x_{k+1})p_k(x)}{x_{k+1} - x_0} \quad (2.17)$$

is the Lagrange interpolation polynomial that interpolates f_0, \dots, f_k , and f_{k+1} at x_0, \dots, x_k , and x_{k+1} .

Proof. Since $p_k(x_i) = q_k(x_i)$ for $i = 1, \dots, k$, we have by a direct calculation by (2.17) that $r_{k+1}(x_i) = f_i$ for $i = 1, \dots, k$. Also by (2.17) we have $r_{k+1}(x_0) = p_k(x_0) = f_0$ and $r_{k+1}(x_{k+1}) = q_k(x_{k+1}) = f_{k+1}$. Therefore, $r_{k+1} \in \mathcal{P}_{k+1}$ is the Lagrange interpolation polynomial that interpolates f_0, \dots, f_{k+1} at x_0, \dots, x_{k+1} . \square

Proof of Theorem 2.8. We prove this theorem by the induction on k , the number of interpolation points. For $k = 0$, clearly $p_0(x) = d_0 = f_0$ is the Lagrange interpolation polynomial that interpolates f_0 at x_0 . Fix an integer k with $0 \leq k \leq n - 1$ and assume that

$$p_j(x) = f[x_0] + f[x_0, x_1](x - x_0) + \dots + f[x_0, \dots, x_j](x - x_0) \dots (x - x_{j-1})$$

interpolates f_0, \dots, f_j at x_0, \dots, x_j for each $j = 0, \dots, k$. We need to show that (2.16) holds true with k replaced by $k + 1$.

Step 1. By the assumption of induction, the polynomial

$$p_k(x) = f[x_0] + f[x_0, x_1](x - x_0) + \dots + f[x_0, \dots, x_k](x - x_0) \dots (x - x_{k-1}) \quad (2.18)$$

is the Lagrange interpolation polynomial that interpolates f_0, \dots, f_k at x_0, \dots, x_k . Let d_{k+1} be given as in (2.15) and

$$p_{k+1}(x) = p_k(x) + d_{k+1}(x - x_0) \dots (x - x_k).$$

Clearly, $p_{k+1}(x_i) = p_k(x_i) = f_i$ for $i = 0, \dots, k$, and by (2.15),

$$p_{k+1}(x_{k+1}) = p_k(x_{k+1}) + d_{k+1}(x_{k+1} - x_0) \dots (x_{k+1} - x_k) = f_{k+1}.$$

Thus, $p_{k+1} \in \mathcal{P}_{k+1}$ is the Lagrange interpolation polynomial that interpolates f_0, \dots, f_k , and f_{k+1} at x_0, \dots, x_k , and x_{k+1} .

Step 2. By the assumption of induction, the polynomial

$$q_k(x) = f[x_1] + f[x_1, x_2](x - x_1) + \cdots + f[x_1, \dots, x_{k+1}](x - x_1) \cdots (x - x_k)$$

is the Lagrange interpolation polynomial that interpolates f_1, \dots, f_{k+1} at x_1, \dots, x_{k+1} . Therefore, it follows from Lemma 2.9 that the polynomial $r_{k+1} \in \mathcal{P}_{k+1}$ defined in (2.17), where p_k is given in (2.18), is the Lagrange interpolation polynomial that interpolates f_0, \dots, f_k , and f_{k+1} at x_0, \dots, x_k , and x_{k+1} . Hence, by the uniqueness of Lagrange interpolation, $r_{k+1} = p_{k+1}$.

Step 3. Comparing the leading coefficients of p_{k+1} and r_{k+1} , we obtain

$$d_{k+1} = \frac{f[x_1, \dots, x_{k+1}] - f[x_0, \dots, x_k]}{x_{k+1} - x_0}.$$

Therefore, (2.16) holds true with k replaced by $k + 1$. □

Theorem 2.10. *Let $p_n \in \mathcal{P}_n$ be the Lagrange interpolation polynomial of a given function $f : [a, b] \rightarrow \mathbb{R}$ at $x_0, \dots, x_n \in [a, b]$. Then for any $x \in [a, b]$ with $x \neq x_i$ ($i = 0, \dots, n$)*

$$f(x) - p_n(x) = f[x_0, \dots, x_n, x](x - x_0) \cdots (x - x_n).$$

Proof. Let $p_{n+1} \in \mathcal{P}_{n+1}$ be the Lagrange interpolation polynomial of f at the $n + 2$ points x_0, \dots, x_n, x . Then by Newton's formula

$$p_{n+1}(t) = p_n(t) + f[x_0, \dots, x_n, x](t - x_0) \cdots (t - x_n)$$

Setting $t = x$, we obtain

$$f(x) = p_{n+1}(x) = p_n(x) + f[x_0, \dots, x_n, x](x - x_0) \cdots (x - x_n),$$

completing the proof. □

Proposition 2.11 (Properties of divided differences). *Let x_0, \dots, x_n be $n + 1$ distinct points in $[a, b]$.*

(1) *Linearity.* For any functions $f, g : \{x_0, \dots, x_n\} \rightarrow \mathbb{R}$ and any $\alpha, \beta \in \mathbb{R}$,

$$(\alpha f + \beta g)[x_0, \dots, x_n] = \alpha f[x_0, \dots, x_n] + \beta g[x_0, \dots, x_n].$$

(2) *Symmetry.* For any function $f : \{x_0, \dots, x_n\} \rightarrow \mathbb{R}$ and any permutation $(i_0 \dots i_n)$ of $(0 \dots n)$,

$$f[x_0, \dots, x_n] = f[x_{i_0}, \dots, x_{i_n}]. \tag{2.19}$$

(3) For any function $f : \{x_0, \dots, x_n\} \rightarrow \mathbb{R}$,

$$f[x_0, \dots, x_n] = \sum_{i=0}^n \frac{f_i}{(x_i - x_0) \cdots (x_i - x_{i-1})(x_i - x_{i+1}) \cdots (x_i - x_n)}. \quad (2.20)$$

(4) If $f \in C^n[a, b]$, then there exists $\xi \in [a, b]$ such that

$$f[x_0, \dots, x_n] = \frac{f^{(n)}(\xi)}{n!}. \quad (2.21)$$

(5) If $f(x) = x^m$ for some integer $m \geq 0$, then

$$f[x_0, \dots, x_n] = \begin{cases} 0 & \text{if } m < n, \\ 1 & \text{if } m = n, \\ \text{linear combinations of } x_0^{k_0} \cdots x_n^{k_n} \text{ with } \sum_{i=0}^n k_i = m - n & \text{if } m > n. \end{cases}$$

Proof. (1) This follows from the definition of divided differences and an argument by induction.

(2) By the uniqueness of the Lagrange interpolation, the Lagrange interpolation polynomial that interpolates f at x_{i_0}, \dots, x_{i_n} is the same as that interpolates f at x_0, \dots, x_n . By Newton's formula (2.16), the leading coefficients in these two polynomials are exactly the right-hand side and left-hand side of (2.19), respectively. Thus, they must be the same.

(3) The left-hand side and right-hand side of (2.20) are the leading coefficients in Newton's formula (2.16) and Lagrange's formula (2.5), respectively, of the unique Lagrange interpolation polynomial that interpolates f at x_0, \dots, x_n .

(4) This is obviously true for $n = 0$. Assume $n \geq 1$. Let $p_{n-1} \in \mathcal{P}_{n-1}$ and $p_n \in \mathcal{P}_n$ be the Lagrange interpolation polynomials that interpolate f_0, \dots, f_{n-1} at x_0, \dots, x_{n-1} and f_0, \dots, f_n at x_0, \dots, x_n , respectively. It follows from Newton's formula (2.16) that

$$p_n(x) = p_{n-1}(x) + f[x_0, \dots, x_n](x - x_0) \cdots (x - x_{n-1}).$$

Hence, replacing x by x_n , we obtain

$$f(x_n) = p_n(x_n) = p_{n-1}(x_n) + f[x_0, \dots, x_n](x_n - x_0) \cdots (x_n - x_{n-1}).$$

On the other hand, by Theorem 2.6 on the remainder of Lagrange interpolation, we have

$$f(x_n) - p_{n-1}(x_n) = \frac{f^{(n)}(\xi)}{n!}(x_n - x_0) \cdots (x_n - x_{n-1})$$

for some $\xi \in [a, b]$. The above two equations imply (2.21).

(5) By (2.21), we need only to consider the case that $m > n$. We use the argument by induction. For $n = 0$, the statement is clearly true. Assume that for any $n \geq 0$ the statement is true, i.e.,

$$f[x_0, \dots, x_n] = \sum_{k_0 + \dots + k_n = m-n} \alpha_{k_0 \dots k_n} x_0^{k_0} \dots x_n^{k_n}, \quad (2.22)$$

where α_{k_0, \dots, k_n} are constants independent of x_0, \dots, x_n . Assume $m > n + 1$. We have by the definition of divided differences, the symmetry property (2.19), and the assumption (2.22) that

$$\begin{aligned} & f[x_0, \dots, x_{n+1}] \\ &= \frac{f[x_1, \dots, x_{n+1}] - f[x_0, \dots, x_n]}{x_{n+1} - x_0} \\ &= \frac{f[x_{n+1}, x_1, \dots, x_n] - f[x_0, x_1, \dots, x_n]}{x_{n+1} - x_0} \\ &= \frac{\sum_{k_0 + \dots + k_n = m-n} \alpha_{k_0 \dots k_n} x_{n+1}^{k_0} x_1^{k_1} \dots x_n^{k_n} - \sum_{k_0 + \dots + k_n = m-n} \alpha_{k_0 \dots k_n} x_0^{k_0} x_1^{k_1} \dots x_n^{k_n}}{x_{n+1} - x_0} \\ &= \sum_{\substack{k_0 + \dots + k_n = m-n \\ k_0 \geq 1}} \alpha_{k_0 \dots k_n} x_1^{k_1} \dots x_n^{k_n} \left(\frac{x_{n+1}^{k_0} - x_0^{k_0}}{x_{n+1} - x_0} \right) \\ &= \sum_{\substack{k_0 + \dots + k_n = m-n \\ k_0 \geq 1}} \alpha_{k_0 \dots k_n} x_1^{k_1} \dots x_n^{k_n} \sum_{k_{n+1}=0}^{k_0-1} x_0^{k_0-1-k_{n+1}} x_{n+1}^{k_{n+1}} \\ &= \sum_{k'_0 + k_1 + \dots + k_{n+1} = m-(n+1)} \alpha_{k'_0 + k_{n+1} + 1, k_2, \dots, k_n} x_0^{k'_0} x_1^{k_1} \dots x_n^{k_n} x_{n+1}^{k_{n+1}}, \end{aligned}$$

where in the last step $k'_0 = k_0 - 1 - k_{n+1}$. This proves that the statement is true for $n + 1$. \square

For each $n \geq 1$, we denote by τ_n the unit simplex in \mathbb{R}^n :

$$\tau_n = \left\{ (t_1, \dots, t_n) \in \mathbb{R}^n : t_i \geq 0 \ (1 \leq i \leq n) \text{ and } \sum_{i=1}^n t_i \leq 1 \right\}.$$

Theorem 2.12 (The Hermite-Genocchi formula). *Let $n \geq 1$ be an integer and $x_0, \dots, x_n \in [0, 1]$ be distinct. We have for any $f \in C^n[0, 1]$ that*

$$f[x_0, \dots, x_n] = \int_{\tau_n} f^{(n)} \left(x_0 + \sum_{j=1}^n t_j (x_j - x_0) \right) dt. \quad (2.23)$$

Proof. We use the argument of induction. The statement is true for $n = 1$, since

$$\int_0^1 f'(x_0 + t_1(x_1 - x_0)) dt_1 = \frac{f(x_1) - f(x_0)}{x_1 - x_0} = f[x_0, x_1].$$

Suppose (2.23) holds true for $n \geq 1$. Consider the case of $n + 1$. We have

$$\begin{aligned} & \int_{\tau_{n+1}} f^{(n+1)} \left(x_0 + \sum_{j=1}^{n+1} t_j(x_j - x_0) \right) dt_1 \cdots dt_{n+1} \\ &= \int_{\tau_n} \left[\int_0^{1-\sum_{j=1}^n t_j} f^{(n+1)} \left(x_0 + \sum_{j=1}^{n+1} t_j(x_j - x_0) \right) dt_{n+1} \right] dt_1 \cdots dt_n \\ &= \frac{1}{x_{n+1} - x_0} \int_{\tau_n} f^{(n)} \left(x_{n+1} + \sum_{j=1}^n t_j(x_j - x_{n+1}) \right) dt_1 \cdots dt_n \\ &\quad - \frac{1}{x_{n+1} - x_0} \int_{\tau_n} f^{(n)} \left(x_0 + \sum_{j=1}^n t_j(x_j - x_0) \right) dt_1 \cdots dt_n \\ &= \frac{f[x_{n+1}, x_1, \dots, x_n] - f[x_0, \dots, x_n]}{x_{n+1} - x_0} \\ &= \frac{f[x_1, \dots, x_{n+1}] - f[x_0, \dots, x_n]}{x_{n+1} - x_0}. \end{aligned}$$

Thus, the statement is true for $n + 1$. This completes the proof. \square

2.3 Peano Kernal and the Remainder Theorem

Let x_0, \dots, x_n be $n + 1$ distinct points in $[a, b]$. Let $L_n : C[a, b] \rightarrow \mathcal{P}_n$ be the Lagrange interpolator associated with x_0, \dots, x_n . We study the error of the Lagrange interpolation $f - L_n f$ for $f \in C[a, b]$ that is not necessary smooth enough, e.g., $f \notin C^{n+1}[a, b]$.

To this end, let us introduce for any integer $k \geq 1$ the function space $W^{k,1}(a, b)$ that consists of all functions $f \in C^{k-1}[a, b]$ such that $f^{(k-1)}$ are absolutely continuous on $[a, b]$. If $f \in W^{k,1}(a, b)$, then $f^{(k)}$ exists as an integrable function on $[a, b]$. Clearly, $C^k[a, b] \subset W^{k,1}(a, b)$.

Let $f \in W^{m+1,1}(a, b)$ with $0 \leq m \leq n$. By the Taylor expansion,

$$\begin{aligned} f(x) &= f(a) + f'(a)(x - a) + \cdots + \frac{f^{(m)}(a)}{m!}(x - a)^m + \frac{1}{m!} \int_a^x (x - t)^m f^{(m+1)}(t) dt \\ &= Q_m(x) + R_m(x), \end{aligned}$$

where

$$Q_m(x) = f(a) + f'(a)(x - a) + \cdots + \frac{f^{(m)}(a)}{m!}(x - a)^m$$

is a polynomial in \mathcal{P}_m and

$$R_m(x) = \frac{1}{m!} \int_a^x (x-t)^m f^{(m+1)}(t) dt$$

is the remainder. Since $Q_m \in \mathcal{P}_m \subseteq \mathcal{P}_n$, $L_n Q_m = Q_m$. Therefore,

$$f - (L_n f) = Q_m + R_m - [L_n Q_m + L_n R_m] = R_m - L_n R_m. \quad (2.24)$$

Let $l_0, \dots, l_n \in \mathcal{P}_n$ be the Lagrange basis polynomials associated with x_0, \dots, x_n . By (2.24) and the Lagrange formula (2.5), we have

$$\begin{aligned} f(x) - (L_n f)(x) &= R_m(x) - (L_n R_m)(x) \\ &= \frac{1}{m!} \int_a^x (x-t)^m f^{(m+1)}(t) dt - \sum_{k=0}^n \left[\frac{1}{m!} \int_a^{x_k} (x_k-t)^m f^{(m+1)}(t) dt \right] l_k(x) \\ &= \frac{1}{m!} \left[\int_a^b (x-t)_+^m f^{(m+1)}(t) dt - \sum_{k=0}^n \left(\int_a^b (x_k-t)_+^m f^{(m+1)}(t) dt \right) l_k(x) \right] \\ &= \int_a^b \frac{1}{m!} \left[(x-t)_+^m - \sum_{k=0}^n (x_k-t)_+^m l_k(x) \right] f^{(m+1)}(t) dt, \end{aligned} \quad (2.25)$$

where

$$c_+ = \begin{cases} c & \text{if } c \geq 0, \\ 0 & \text{if } c < 0. \end{cases}$$

We define

$$K_m(x, t) = \frac{1}{m!} \left[(x-t)_+^m - \sum_{k=0}^n (x_k-t)_+^m l_k(x) \right] = \frac{1}{m!} E_n((\cdot - t)_+^m)(x), \quad (2.26)$$

where $E_n(g) = g - L_n g$ is the error of the Lagrange interpolation for $g \in C[a, b]$. We shall call $K_m : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ the m th *Peano kernel* associated with the interpolation points x_0, \dots, x_n .

We have in fact proved the following:

Theorem 2.13 (The Peano Remainder Theorem for the Lagrange interpolation).

Let m be an integer with $0 \leq m \leq n$. Then for any $f \in W^{m+1,1}(a, b)$

$$f(x) - (L_n f)(x) = \int_a^b K_m(x, t) f^{(m+1)}(t) dt \quad \forall x \in [a, b]. \quad (2.27)$$

To estimate the interpolation error using the Peano kernel representation (2.27), we introduce the function space $W^{k,\infty}(a, b)$ that consists of all the functions $f \in C^{k-1}[a, b]$

such that $f^{(k-1)}$ are Lipschitz continuous on $[a, b]$. If $f \in W^{k,\infty}(a, b)$, then $f^{(k)}$ exists as an integrable and bounded function on $[a, b]$. It can be proved that for any integer $k \geq 1$

$$C^k[a, b] \subsetneq W^{k,\infty}(a, b) \subsetneq W^{k,1}(a, b) \subsetneq C[a, b].$$

We denote for a measurable and bounded function $g : [a, b] \rightarrow \mathbb{R}$

$$\|g\|_{L^\infty(a,b)} = \sup_{a \leq x \leq b} |g(x)|.$$

Theorem 2.14 (Error estimates for the Lagrange interpolation error). *Let m be an integer with $0 \leq m \leq n$. Then for any $f \in W^{m+1,\infty}(a, b)$,*

$$|f(x) - (L_n f)(x)| \leq \left[\int_a^b |K_m(x, t)| dt \right] \|f^{(m+1)}\|_{L^\infty(a,b)} \quad \forall x \in [a, b]. \quad (2.28)$$

Moreover, for any $x \in [a, b]$, there exists $f_0 \in W^{m+1,\infty}(a, b)$ such that

$$|f_0(x) - (L_n f_0)(x)| = \left[\int_a^b |K_m(x, t)| dt \right] \|f_0^{(m+1)}\|_{L^\infty(a,b)}. \quad (2.29)$$

Proof. Let $f \in W^{m+1,\infty}(a, b)$. By (2.25) and (2.26), we have

$$|f(x) - (L_n f)(x)| \leq \int_a^b |K_m(x, t)| |f^{(m+1)}(t)| dt \leq \left[\int_a^b |K_m(x, t)| dt \right] \|f^{(m+1)}\|_{L^\infty(a,b)},$$

implying (2.28). Fix $x \in [a, b]$. Define $g_m(t) = \text{sign } K_m(x, t)$ and

$$f_0(x) = \underbrace{\int_a^t \dots \int_a^t}_{m+1 \text{ times}} g_m(t) \underbrace{dt \dots dt}_{m+1 \text{ times}}.$$

Then, $f_0 \in W^{m+1,\infty}(a, b)$ and $\|f_0^{(m+1)}\|_{L^\infty(a,b)} = \|g_m\|_{L^\infty(a,b)} = 1$. By (2.25) and (2.26), we obtain

$$\begin{aligned} |f_0(x) - (L_n f_0)(x)| &= \left| \int_a^b K_m(x, t) f_0^{(m+1)}(t) dt \right| = \left| \int_a^b K_m(x, t) g_m(t) dt \right| \\ &= \int_a^b |K_m(x, t)| dt = \left[\int_a^b |K_m(x, t)| dt \right] \|f_0^{(m+1)}\|_{L^\infty(a,b)}, \end{aligned}$$

leading to (2.29). □

Theorem 2.15. (1) *We have*

$$\int_a^b K_n(x, t) dt = \frac{1}{(n+1)!} \prod_{k=0}^n (x - x_k) \quad \forall x \in [a, b]. \quad (2.30)$$

(2) If $a = \min_{0 \leq k \leq n} x_k$ and $b = \max_{0 \leq k \leq n} x_k$, then

$$\int_a^b |K_n(x, t)| dt = \frac{1}{(n+1)!} \prod_{k=0}^n |x - x_k| \quad \forall x \in [a, b]. \quad (2.31)$$

Proof. (1) Let $Q_{n+1}(t) = t^{n+1}/(n+1)!$. By Theorem 2.6, we have for any $x \in [a, b]$ that

$$Q_{n+1}(x) - (L_n Q_{n+1})(x) = \frac{1}{(n+1)!} \prod_{k=0}^n (x - x_k).$$

On the other hand, we have by Theorem 2.13 that

$$Q_{n+1}(x) - (L_n Q_{n+1})(x) = \int_a^b K_n(x, t) Q_{n+1}^{(n+1)}(t) dt = \int_a^b K_n(x, t) dt.$$

Thus, (2.30) holds true.

(2) Let $x \in [a, b]$. By the lemma below, $K_n(x, \cdot)$ does not change its sign in (a, b) . Thus,

$$\left| \int_a^b K_n(x, t) dt \right| = \int_a^b |K_n(x, t)| dt.$$

This and (2.30) imply (2.31). □

Lemma 2.16. Assume $a = \min_{0 \leq k \leq n} x_k$ and $b = \max_{0 \leq k \leq n} x_k$. Assume $0 \leq m \leq n$ and $x \in [a, b]$. Then, $K_m(x, \cdot)$ changes its sign in (a, b) exactly $n - m$ times. In particular, $K_n(x, \cdot)$ does not change its sign in (a, b) .

Proof. The statement is trivially true for the case $n = 0$, since $K_0(x, \cdot) = 0$ by (2.26). Assume $n \geq 1$. We divide our proof into three steps.

Step 1. If $0 \leq m \leq n - 1$, then the function $K_m(x, \cdot) \in C[a, b]$ changes its sign in (a, b) at least once. This follows from an application of (2.27) to $f = Q_{m+1} \in \mathcal{P}_{m+1} \subseteq \mathcal{P}_n$ with $Q_{m+1}(t) = t^{m+1}/(m+1)!$ for which $L_n Q_{m+1} = Q_{m+1}$:

$$0 = Q_{m+1}(x) - (L_n Q_{m+1})(x) = \int_a^b K_m(x, t) Q_{m+1}^{(m+1)}(t) dt = \int_a^b K_m(x, t) dt.$$

Step 2. Let $0 \leq m \leq n - 1$. If $K_m(x, \cdot)$ changes sign in (a, b) exactly k times, then $K_{m+1}(x, \cdot)$ changes its sign in (a, b) at most $k - 1$ times. This follows from the fact that $(d/dt)K_{m+1}(x, t) = -K_m(x, t)$ and $K_{m+1}(x, a) = K_{m+1}(x, b) = 0$, and an application of Rolle's Theorem.

Step 3. For $0 \leq m \leq n$, $K_m(x, \cdot)$ changes its sign in (a, b) exactly $n - m$ times. In particular, $K_n(x, \cdot)$ does not change its sign in (a, b) .

The function $K_0(x, \cdot)$ is a piecewise constant. It has jumps at x_0, \dots, x_n, x . By the assumption of the lemma, two of these points are a and b . Therefore, $K_0(x, \cdot)$ can change

its sign in (a, b) at most n times. By Step 2 and an argument of induction, $K_m(x, \cdot)$ can change its sign in (a, b) at most $n - m$ times. If for some m with $0 \leq m \leq n - 1$, $K_m(x, \cdot)$ changes its sign in (a, b) less than $n - m$ times, then by Step 2 and induction, $K_{n-1}(x, \cdot)$ changes its sign in (a, b) less than $n - (n - 1) = 1$ times. By step 1, this is impossible. Thus, for each m with $0 \leq m \leq n - 1$, $K_m(x, \cdot)$ changes its sign in (a, b) exactly $n - m$ times. By Step 2, $K_n(x, \cdot)$ does not change its sign in (a, b) . \square

2.4 Hermite Interpolation and Divided Differences with Repeated Points

Theorem 2.17. *Let x_1, \dots, x_n be n distinct points in $[a, b]$. Let y_1, \dots, y_n and y'_1, \dots, y'_n be $2n$ real numbers. Then there exists a unique $p \in \mathcal{P}_{2n-1}$ such that*

$$p(x_k) = y_k, \quad p'(x_k) = y'_k, \quad k = 1, \dots, n. \quad (2.32)$$

Moreover, p is given by

$$p(x) = \sum_{k=1}^n y_k [1 - 2l'_k(x_k)(x - x_k)][l_k(x)]^2 + \sum_{k=1}^n y'_k (x - x_k)[l_k(x)]^2, \quad (2.33)$$

where $l_k \in \mathcal{P}_{n-1}$ ($k = 1, \dots, n$) are the Lagrange basis polynomials associated with x_1, \dots, x_n .

Proof. Define for each integer k with $1 \leq k \leq n$

$$\phi_k(x) = [1 - 2l'_k(x_k)(x - x_k)][l_k(x)]^2, \quad (2.34)$$

$$\psi_k(x) = (x - x_k)[l_k(x)]^2. \quad (2.35)$$

Clearly, all ϕ_k, ψ_k are polynomials of degree $2n - 1$. Moreover,

$$\begin{aligned} \phi'_k(x) &= -2l'_k(x_k)[l_k(x)]^2 + [1 - 2l'_k(x_k)(x - x_k)]2l_k(x)l'_k(x), \\ \psi'_k(x) &= [l_k(x)]^2 + 2(x - x_k)l_k(x)l'_k(x). \end{aligned}$$

Therefore,

$$\phi_k(x_j) = \delta_{kj}, \quad \phi'_k(x_j) = 0, \quad \psi_k(x_j) = 0, \quad \psi'_k(x_j) = \delta_{kj}, \quad j, k = 1, \dots, n. \quad (2.36)$$

By the definition of ϕ_k and ψ_k (cf. (2.34) and (2.35)), the polynomial p defined in (2.33) is

$$p = \sum_{k=1}^n y_k \phi_k + \sum_{k=1}^n y'_k \psi_k.$$

Clearly, $p \in \mathcal{P}_{2n-1}$. Moreover, (2.32) and (2.36) imply (2.32).

To prove the uniqueness, we assume $q \in \mathcal{P}_{2n-1}$ also satisfies that $q(x_k) = y_k$ and $q'(x_k) = y'_k$ for $k = 1, \dots, n$. Then $r := p - q \in \mathcal{P}_{2n-1}$ and $r(x_k) = r'(x_k) = 0$ for all $k = 1, \dots, n$. Therefore, r has $2n$ roots, counting the multiplicity of each root. Hence $r = 0$ and $q = p$. \square

We call p the *Hermite interpolation polynomial*, or *Hermite interpolant*, of y_k, y'_k at x_1, \dots, x_n . If $y_k = f(x_k)$ and $y'_k = f'(x_k)$ ($k = 1, \dots, n$) for some $f \in C^1[a, b]$, then p is called the Hermite interpolation polynomial (or Hermite interpolant) of f at x_1, \dots, x_n .

Example. Consider $n = 2$, $x_1 = a$, and $x_2 = b$. We have

$$l_1(x) = \frac{x-b}{a-b} \quad \text{and} \quad l_2(x) = \frac{x-a}{b-a}.$$

Therefore,

$$\begin{aligned} \phi_1(x) &= \left[1 + \frac{2(x-a)}{b-a} \right] \left(\frac{x-b}{b-a} \right)^2, \\ \phi_2(x) &= \left[1 - \frac{2(x-b)}{b-a} \right] \left(\frac{x-a}{b-a} \right)^2, \\ \psi_1(x) &= \frac{(x-a)(x-b)^2}{(b-a)^2}, \\ \psi_2(x) &= \frac{(x-a)^2(x-b)}{(b-a)^2}. \end{aligned}$$

The Hermite interpolation polynomial $p_3 \in \mathcal{P}_3$ that interpolates y_1, y_2 and y'_1, y'_2 at $x_1 = a, x_2 = b$ is

$$\begin{aligned} p_3(x) &= y_1\phi_1(x) + y_2\phi_2(x) + y'_1\psi_1(x) + y'_2\psi_2(x) \\ &= y_1 \left[1 + \frac{2(x-a)}{b-a} \right] \left(\frac{x-b}{b-a} \right)^2 + y_2 \left[1 - \frac{2(x-b)}{b-a} \right] \left(\frac{x-a}{b-a} \right)^2 \\ &\quad + y'_1 \frac{(x-a)(x-b)^2}{(b-a)^2} + y'_2 \frac{(x-a)^2(x-b)}{(b-a)^2}. \end{aligned}$$

Theorem 2.18 (The remainder of Hermite interpolation). Let $f \in C^{2n}[a, b]$. Let $H_{2n-1}f \in \mathcal{P}_{2n-1}$ be the Hermite interpolation polynomial of f at x_1, \dots, x_n . Then for any $x \in [a, b]$ there exists $\xi = \xi(x) \in [a, b]$ such that

$$f(x) - (H_{2n-1}f)(x) = \frac{f^{(2n)}(\xi(x))}{(2n)!} (x-x_1)^2 \cdots (x-x_n)^2. \quad (2.37)$$

Proof. Let $x \in [a, b]$. If $x = x_k$ for some i then $\xi(x) \in [a, b]$ can be any number. So, let us assume $x \neq x_k$ ($k = 1, \dots, n$). Let

$$g(t) = f(t) - (H_{2n-1}f)(t) - \lambda(t-x_1)^2 \cdots (t-x_n)^2,$$

where $\lambda \in \mathbb{R}$ is so chosen that $g(x) = 0$, i.e.,

$$\lambda = \frac{f(x) - (H_{2n-1}f)(x)}{(x-x_1)^2 \cdots (x-x_n)^2}.$$

Notice that $g = 0$ at $n + 1$ distinct points x, x_1, \dots, x_n . Thus it follows from Rolle's Theorem that there exist $\xi_1, \dots, \xi_n \in [a, b]$ with $\xi_j \neq x$ and $\xi_j \neq x_k$ ($k = 1, \dots, n$) for each j ($1 \leq j \leq n$). Notice also that $g' = 0$ at x_1, \dots, x_n . Therefore, $g' = 0$ at $2n$ points $x, x_1, \dots, x_n, \xi_1, \dots, \xi_n$. Applying Rolle's Theorem repeatedly, we conclude that $g^{(2n)} = 0$ at some point $\xi = \xi(x) \in [a, b]$. This, together with the fact that

$$g^{(2n)}(t) = f^{(2n)}(t) - \lambda(2n)!$$

and the definition of λ , leads to (2.37). □

2.5 Convergence of Interpolation Polynomials

Let $n \geq 0$ be an integer, $x_0^{(n)}, \dots, x_n^{(n)}$ be $n + 1$ distinct points in $[a, b]$, and $L_n : C[a, b] \rightarrow \mathcal{P}_n$ the associated Lagrange interpolator. Does $\{L_n f(x)\}_{n=0}^\infty$ converge to $f(x)$ for any $f \in C[a, b]$ and any $x \in [a, b]$? It turns out there are many negative results.

Runge's example. Consider $[a, b] = [-5, 5]$ and $x_k^{(n)}$ evenly spaced in $[-5, 5]$, i.e., $x_k^{(n)} = -5 + 10k/n$ ($k = 0, \dots, n; n = 1, \dots$). For $f(x) = 1/(1 + x^2)$, Runge (year?) proved that there exists $\kappa \approx 3.63338$ such that

$$\lim_{n \rightarrow \infty} (L_n f)(x) = f(x) \quad \text{if and only if} \quad |x| < \kappa.$$

See more details in [5] (Section 3.4 of Chapter 6).

Bernstein (19??). For $[a, b] = [-1, 1]$, evenly spaced interpolation points $x_k^{(n)} \in [-1, 1]$ ($k = 0, \dots, n; n = 1, \dots$), and the function $f(x) = |x|$, Bernstein (19??) proved that

$$\lim_{n \rightarrow \infty} L_n f(x) = f(x) \quad \text{if and only if} \quad x \in \{0, 1, -1\}.$$

Faber (1914). In 1914, Faber proved the following: For any given sequence of interpolation points $x_k^{(n)} \in [a, b]$ ($k = 0, \dots, n; n = 0, \dots$), there exists $f \in C[a, b]$ such that $\|L_n f - f\|_{C[a, b]} \not\rightarrow 0$.

Bernstein (1931). In 1931, Bernstein proved the following result: For any given sequence of interpolation points $x_k^{(n)} \in [a, b]$ ($k = 0, \dots, n; n = 0, \dots$), there exist $f \in C[a, b]$ and $x \in [a, b]$ such that $(L_n f)(x) \not\rightarrow f(x)$.

Erdős and Vértesi (1980). In 1980, Erdős and Vértesi proved the following striking negative result: For any given sequence of interpolation points $x_k^{(n)} \in [a, b]$ ($k = 0, \dots, n; n = 0, \dots$), there exist $f \in C[a, b]$ such that $(L_n f)(x) \not\rightarrow f(x)$ for almost all $x \in [a, b]$.

There are also some positive results.

Theorem 2.19. For any sequence of Lagrange interpolators $L_n : C[a, b] \rightarrow \mathcal{P}_n$ ($n = 0, \dots$), $\|L_n f - f\|_{C[a, b]} \rightarrow 0$ for any $f \in C[a, b]$ that is the restriction onto $[a, b]$ of an entire function.

Proof. Let □

Theorem 2.20. For any $f \in C[a, b]$, there exist $n + 1$ distinct points $x_0^{(n)}, \dots, x_n^{(n)} \in [a, b]$ for each $n \geq 0$ such that

$$\lim_{n \rightarrow \infty} \|f - L_n f\|_{C[a, b]} = 0,$$

where $L_n : C[a, b] \rightarrow \mathcal{P}_n$ is the Lagrange interpolator associated with $x_0^{(n)}, \dots, x_n^{(n)}$ ($n = 0, \dots$).

Proof. Fix $f \in C[a, b]$. If $f \in \mathcal{P}$, then we can choose $x_0^{(n)}, \dots, x_n^{(n)} \in [a, b]$ to be any $n + 1$ distinct points for each $n \geq 0$. Clearly, $L_n f = f$ for n sufficiently large.

Assume $f \notin \mathcal{P}$. Let $n \geq 0$ be an integer. Let $p_n \in \mathcal{P}_n$ be the best uniform approximation of f in \mathcal{P}_n . Then it follows from the Chebyshev Alternation Theorem that there exist $n + 1$ distinct points $x_k^{(n)}$ ($k = 0, \dots, n$) such that

$$f(x_k^{(n)}) - p_n(x_k^{(n)}) = 0, \quad k = 0, \dots, n.$$

Therefore, $p_n = L_n f$ is the Lagrange interpolation polynomial of f at $x_0^{(n)}, \dots, x_n^{(n)}$. Consequently, we have by Proposition 1.4 that

$$\|f - L_n f\|_{C[a, b]} = \|f - p_n\|_{C[a, b]} = \min_{q \in \mathcal{P}_n} \|f - q\|_{C[a, b]} \rightarrow 0 \quad \text{as } n \rightarrow \infty,$$

proving the theorem. □

Theorem 2.21. Let $L_{n-1} : C[-1, 1] \rightarrow \mathcal{P}_{n-1}$ be the Lagrange interpolator associated with the n roots of the Chebyshev polynomial T_n ($n = 1, \dots$). Then for any $f \in C^2[-1, 1]$

$$\|L_n f - f\|_{C[-1, 1]} = O\left(\frac{1}{\sqrt{n}}\right) \quad \text{as } n \rightarrow \infty.$$

Proof. See the textbook p. 236. □

Theorem 2.22. (B.H. KpбиπoB, 1956) $[a, -b] = [-1, 1]$.

$x_i^{(n)}$: zeros of Chebyshev polynomials.

1. $f \in C[a, b]$ is absolute continuous on $[a, b]$.
 $\Rightarrow \|L_n f - f\| \rightarrow 0$.
2. f has a total variation on $[-1, 1]$.
 $\Rightarrow (L_n f)(x) \rightarrow f(x)$ at any continuity $x \in [a, b]$.

Theorem 2.23 (Erdős-Turán (1937)). Let $x_1^{(n)}, \dots, x_n^{(n)}$ be the n distinct roots of orthogonal polynomials Q_n ($n = 1, \dots$) in $L_\rho^2(a, b)$. For each $n \geq 1$, let $L_{n-1} : C[a, b] \rightarrow \mathcal{P}_{n-1}$ be the Lagrange interpolator associated with $x_1^{(n)}, \dots, x_n^{(n)}$. Then

$$\lim_{n \rightarrow \infty} \int_a^b \rho(x) [f(x) - (L_{n-1}f)(x)]^2 dx = 0 \quad \forall f \in C[a, b]. \quad (2.38)$$

To prove this theorem, we need the following lemma.

Lemma 2.24. Let $x_1^{(n)}, \dots, x_n^{(n)}$ be the n distinct roots of orthogonal polynomials Q_n ($n = 1, \dots$) in $L_\rho^2(a, b)$. For each $n \geq 1$, let $l_1^{(n)}, \dots, l_n^{(n)}$ be the Lagrange basis polynomials associated with $x_1^{(n)}, \dots, x_n^{(n)}$. Then

$$\int_a^b \rho(x) l_j^{(n)}(x) l_k^{(n)}(x) dx = 0 \quad \text{if } 1 \leq j, k \leq n, \text{ and } j \neq k, \quad (2.39)$$

$$\sum_{k=1}^n \int_a^b \rho(x) [l_k^{(n)}(x)]^2 dx = \int_a^b \rho(x) dx. \quad (2.40)$$

Proof. Without loss of generality, we assume that $n \geq 2$. Fix j, k with $1 \leq j, k \leq n$ and $j \neq k$. The polynomial $(x - x_k^{(n)}) l_k^{(n)}(x)$ in \mathcal{P}_n has n simple roots $x_1^{(n)}, \dots, x_n^{(n)}$. Thus, there exists a constant $\alpha_k^{(n)}$ such that $(x - x_k^{(n)}) l_k^{(n)}(x) = Q_n(x)$ for all x . The polynomial $l_j^{(n)}(x) / (x - x_k^{(n)})$ has degree $n - 2$. Therefore, we have by the orthogonality that

$$\begin{aligned} \int_a^b \rho(x) l_j^{(n)}(x) l_k^{(n)}(x) dx &= \int_a^b \rho(x) \left[\frac{l_j^{(n)}(x)}{x - x_k^{(n)}} \right] (x - x_k^{(n)}) l_k^{(n)}(x) dx \\ &= \alpha_k^{(n)} \int_a^b \rho(x) \left[\frac{l_j^{(n)}(x)}{x - x_k^{(n)}} \right] Q_n(x) dx \\ &= 0, \end{aligned}$$

proving (2.39).

By (2.6) with $p_n(x) = 1$, we have $\sum_{k=1}^n l_k^{(n)}(x) = 1$ identically. Thus, it follows from (2.39) that

$$\begin{aligned} \int_a^b \rho(x) dx &= \int_a^b \rho(x) \left[\sum_{k=1}^n l_k^{(n)}(x) \right]^2 dx \\ &= \sum_{j,k=1}^n \int_a^b \rho(x) l_j^{(n)}(x) l_k^{(n)}(x) dx \end{aligned}$$

$$= \sum_{k=1}^n \int_a^b \rho(x) \left[l_k^{(n)}(x) \right]^2 dx.$$

This is (2.40). □

Proof of Theorem 2.23. Let $n \geq 2$ and let $p_{n-1} \in \mathcal{P}_{n-1}$ be the best uniform approximation of f in \mathcal{P}_{n-1} . We have

$$\begin{aligned} & \int_a^b \rho(x) [f(x) - L_{n-1}f(x)]^2 dx \\ & \leq 2 \int_a^b \rho(x) [f(x) - p_{n-1}(x)]^2 dx + 2 \int_a^b \rho(x) [p_{n-1}(x) - L_{n-1}f(x)]^2 dx. \end{aligned} \quad (2.41)$$

By Proposition 1.4,

$$\int_a^b \rho(x) [f(x) - p_{n-1}(x)]^2 dx \leq \|f - p_{n-1}\|_{C[a,b]}^2 \int_a^b \rho(x) dx \rightarrow 0 \quad \text{as } n \rightarrow \infty. \quad (2.42)$$

It follows from Proposition 2.5, Lemma 2.24, and Proposition 1.4 that

$$\begin{aligned} & \int_a^b \rho(x) [L_{n-1}f(x) - p_{n-1}(x)]^2 dx \\ & = \int_a^b \rho(x) \{ [L_{n-1}(f - p_{n-1})](x) \}^2 dx \\ & = \int_a^b \rho(x) \sum_{j,k=1}^n \left[f(x_j^{(n)}) - p_{n-1}(x_j^{(n)}) \right] \left[f(x_k^{(n)}) - p_{n-1}(x_k^{(n)}) \right] l_j^{(n)}(x) l_k^{(n)}(x) dx \\ & = \sum_{j,k=1}^n \left[f(x_j^{(n)}) - p_{n-1}(x_j^{(n)}) \right] \left[f(x_k^{(n)}) - p_{n-1}(x_k^{(n)}) \right] \int_a^b \rho(x) l_j^{(n)}(x) l_k^{(n)}(x) dx \\ & = \sum_{k=1}^n \left[f(x_k^{(n)}) - p_{n-1}(x_k^{(n)}) \right]^2 \int_a^b \rho(x) \left[l_k^{(n)}(x) \right]^2 dx \\ & \leq \|f - p_{n-1}\|_{C[a,b]}^2 \sum_{k=1}^n \int_a^b \rho(x) \left[l_k^{(n)}(x) \right]^2 dx \\ & = \|f - p_{n-1}\|_{C[a,b]}^2 \int_a^b \rho(x) dx \\ & \rightarrow 0 \quad \text{as } n \rightarrow \infty. \end{aligned}$$

This, together with (2.41) and (2.42), implies (2.38). □

For any integer $n \geq 1$ and any n distinct points $x_1, \dots, x_n \in [a, b]$, we define the *Fajér-Hermite operator* $F_n : C[a, b] \rightarrow \mathcal{P}_{2n-1}$ by

$$F_n f = \sum_{k=1}^n f(x_k) \phi_k(x) \quad \forall f \in C[a, b],$$

where ϕ_k is defined in (2.34).

Theorem 2.25. *Let $F_n : C[-1, 1] \rightarrow \mathcal{P}_{2n-1}$ be the Fajér-Hermite operator associated with the zeros of Chebyshev polynomial T_n . Then*

$$\lim_{n \rightarrow \infty} \|f - F_n f\| = 0 \quad \forall f \in C[-1, 1].$$

Proof. By Bohman-Korovkin Theorem, we need only to show ... □

2.6 Splines

2.7 Trigonometric Interpolation and Fast Fourier Transforms

Exercises

1. Find the polynomial $p \in \mathcal{P}_3$ of the form $p(x) = c_0 + c_1x + x^3$ that interpolates a given function $f \in C[0, 3]$ at $x = 0, 2, 3$.
2. Let $x_0 = 2, x_1 = 3, x_2 = 5, x_3 = 6$ and $y_0 = 5, y_1 = 2, y_2 = 3, y_3 = 4$. Let $p \in \mathcal{P}_3$ be the unique polynomial that interpolates y_j at x_j ($j = 0, 1, 2, 3$). Calculate p by using: (a) Lagrange's formula; and (b) Newton's formula.
3. Let $f(x) = x^4 - x^2 + 17x + 1$. Let $p \in \mathcal{P}_{20}$ interpolates f at $x_j = 2^j$ ($j = 0, \dots, 20$). Compute $p(0)$.
4. Find an approximation of $\sqrt{3}$ with the values of the function $f(x) = 3^x$ at $x_0 = 0, x_1 = 1$, and $x_2 = 2$ using
 - (a) Aitken's iterative linear interpolation method;
 - (b) Neville's iterative linear interpolation method.
5. Let x_0, \dots, x_n be $n + 1$ distinct real numbers. Let $l_j(x)$ be the associated Lagrange basis polynomials. Show that

$$\sum_{j=0}^n (x - x_j)^k l_j(x) = 0 \quad \forall k = 1, \dots, n.$$

6. Let x_0, \dots, x_n be $n + 1$ distinct real numbers. Let $l_j(x)$ be the associated Lagrange basis polynomials. Show for any j with $1 \leq j \leq n$ that

$$\sum_{i=0}^n |l_i(x)| = \left| \sum_{i=0}^{j-1} (-1)^i l_i(x) - \sum_{i=j}^n (-1)^i l_i(x) \right| \quad \forall x \in (x_{j-1}, x_j).$$

7. Recall for $n \geq 1$ that the Chebyshev polynomial $T_n(x)$ has n distinct roots $x_j = \cos \theta_j$ with $\theta_j = (2j - 1)\pi/2n$ ($j = 1, \dots, n$). Denote by $L_{n-1} : C[-1, 1] \rightarrow \mathcal{P}_{n-1}$ the associated Lagrange interpolation operator. Show that

$$(L_{n-1}f)(x) = \frac{1}{n} \sum_{j=1}^n f(x_j) \frac{(-1)^{j-1} \sin \theta_j T_n(x)}{x - x_j} \quad \forall f \in C[-1, 1].$$

8. Let $Q_n \in \mathcal{P}_n$ ($n = 0, 1, \dots$) be orthonormal polynomials in $L^2_\rho[a, b]$. Fix $n \geq 2$. Let x_1, \dots, x_n be the n distinct roots of $Q_n(x)$ in (a, b) , and l_1, \dots, l_n be the associated Lagrange basis polynomials.

- (a) Prove that l_1, \dots, l_n are orthogonal in $L^2_\rho[a, b]$.
 (b) Prove the identity

$$\sum_{j=1}^n \int_a^b \rho(x) [l_j(x)]^2 dx = \int_a^b \rho(x) dx.$$

9. Let x_0, \dots, x_n be $n + 1$ distinct points in $[a, b]$ and $L_n : C[a, b] \rightarrow \mathcal{P}_n$ the corresponding Lagrange interpolation operator. Show that

$$\|L_n f\|_{C[a,b]} \leq \lambda_n \|f\|_{C[a,b]} \quad \forall f \in C[a, b],$$

where

$$\lambda_n = \max_{a \leq x \leq b} \sum_{j=0}^n |l_j(x)|$$

and l_0, \dots, l_n are the Lagrange basis polynomials associated with x_0, \dots, x_n . Show also that there exists $\tilde{f} \in C[a, b]$ depending on n such that

$$\|L_n \tilde{f}\|_{C[a,b]} = \lambda_n \|\tilde{f}\|_{C[a,b]}.$$

10. Let $\{x_j\}_{j=0}^\infty$ be a sequence of equidistant points $x_j = x_0 + jh$ with $h > 0$. Define for each $j \geq 0$

$$\Delta^0 f(x_j) = f(x_j) \quad \text{and} \quad \Delta^k f(x_j) = \Delta^{k-1} f(x_{j+1}) - \Delta^{k-1} f(x_j), \quad k = 1, \dots$$

(a) Let $f \in C^n[x_0, x_n]$. Prove that

$$f[x_0, \dots, x_n] = \frac{1}{n!h^n} \Delta^n f(x_0)$$

and that

$$\Delta^n f(x_0) = h^n f^{(n)}(\xi)$$

for some $\xi \in [x_0, x_n]$.

(b) Let $f \in C^{n+1}[x_0, x_n]$. Let $p_n \in \mathcal{P}_n$ be the unique Lagrange polynomial that interpolates f at x_0, \dots, x_n . Let t be a real number. Show that

$$p_n(x_0 + th) = \frac{\pi_n(t)}{n!} \sum_{j=0}^n (-1)^{n-j} \binom{n}{j} \frac{f(x_j)}{t-j}$$

and that

$$p_n(x_0 + th) = f(x_0) + \frac{\pi_0(t)}{1!} \Delta f(x_0) + \frac{\pi_1(t)}{2!} \Delta^2 f(x_0) + \dots + \frac{\pi_{n-1}(t)}{n!} \Delta^n f(x_0),$$

where

$$\pi_0(t) = t, \quad \text{and} \quad \pi_j(t) = t(t-1) \cdots (t-j), \quad j = 1, \dots, n.$$

Show also that

$$f(x_0 + th) - p_n(x_0 + th) = \pi_n(t) h^{n+1} \frac{f^{(n+1)}(\eta)}{(n+1)!}$$

for some η in an interval containing all x_0, \dots, x_n and $x_0 + th$.

11. Let $N \geq 1$ be an integer, $h = (b-a)/N$, and $x_j = a + jh$, $j = 0, \dots, N$. For any $f \in C[a, b]$, let $I_h f \in C[a, b]$ be such that $I_h f \in \mathcal{P}_1$ on each $[x_{j-1}, x_j]$ ($j = 1, \dots, N$) and $I_h f(x_j) = f(x_j)$ ($j = 0, \dots, N$).

(a) Suppose $f \in C^2[a, b]$. Show that

$$\max_{a \leq x \leq b} |f(x) - (I_h f)(x)| \leq \frac{1}{8} M_2 h^2,$$

where $M_2 = \max_{a \leq x \leq b} |f''(x)|$.

(b) Suppose $f \in C^3[a, b]$. Show that

$$\max_{1 \leq j \leq N} \sup_{a < x < b} |f'(x) - (I_h f)'(x)| \leq \frac{M_2}{2} h + \frac{M_3}{24} h^2,$$

where $M_k = \max_{a \leq x \leq b} |f^{(k)}(x)|$ for $k = 2$ or 3 , and that

$$\max_{1 \leq j \leq N} |f'(m_j) - (I_h f)'(m_j)| \leq \frac{M_3}{24} h^2,$$

where $m_j = (x_{j-1} + x_j)/2$ is the midpoint of the interval $[x_{j-1}, x_j]$, $j = 1, \dots, N$.

12. For each integer $n \geq 0$, let $x_0^{(n)}, \dots, x_n^{(n)}$ be $n + 1$ distinct points in $[a, b]$. Let $L_n : C[a, b] \rightarrow \mathcal{P}_n$ be the associated Lagrange interpolation operator. Let $f \in C^\infty[a, b]$ satisfy for some constant $M > 0$ that

$$\|f^{(k)}\|_{C[a,b]} \leq M \quad \forall k \geq 1.$$

Show that

$$\|f - L_n f\|_{C[a,b]} \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

13. Let $f \in C[a, b]$. Show that, for each integer $n \geq 1$, there exist n distinct points $x_1^{(n)}, \dots, x_n^{(n)}$ such that

$$\|f - L_{n-1} f\|_{C[a,b]} \rightarrow 0 \quad \text{as } n \rightarrow \infty,$$

where $L_{n-1} f \in \mathcal{P}_{n-1}$ is the Lagrange interpolation polynomial of f at $x_1^{(n)}, \dots, x_n^{(n)}$.

14. Let $n \geq 1$ be an integer. Define

$$\pi(t) = \prod_{j=0}^n (t - j) \quad \text{and} \quad l_j(t) = \frac{\pi(t)}{(t - j)\pi'_n(j)} \quad j = 0, \dots, n.$$

Show that

$$|\pi(t)| \leq n! \quad \forall t \in [0, n],$$

that

$$|l_j(t)| \leq \binom{n}{j} \quad \forall t \in [0, n], \quad j = 0, \dots, n,$$

and that

$$\sum_{j=0}^n |l_j(t)| \leq 2^n.$$

15. Let $f \in C[a, b]$. Let $n \geq 1$ be a fixed integer. For each integer $N \geq 1$, let $H_N = (b - a)/N$ and $x_j^{(N)} = a + jH_N$, $j = 0, \dots, N$. Let $p_n^{(N)} \in C[a, b]$ satisfy for each j with $1 \leq j \leq N$ that the restriction of $p_n^{(N)}$ on the subinterval $[x_{j-1}^{(N)}, x_j^{(N)}]$ is the Lagrange interpolation polynomial in \mathcal{P}_n that interpolates f at the $n + 1$ points $x_{j-1}^{(N)} + kH_N/n$, $k = 0, \dots, n$.

(a) Show that

$$\|f - p_n^{(N)}\|_{C[a,b]} \leq 2^n \omega_f(H_N),$$

where ω_f is the modulus of continuity of f , and that

$$\|f - p_n^{(N)}\|_{C[a,b]} \rightarrow 0 \quad \text{as } N \rightarrow \infty.$$

(b) If $f \in C^{n+1}[a, b]$, show that

$$\|f - p_n^{(N)}\|_{C[a,b]} \leq \frac{\|f^{(n+1)}\|_{C[a,b]}}{n+1} \left(\frac{H_N}{n}\right)^{n+1}.$$

16. Let $f(x) = \sin x$, $[a, b] = [0, 1]$, and $n = 1$. Find an integer $N \geq 1$, as small as possible, such that

$$\|f - p_n^{(N)}\|_{C[a,b]} \leq 1.25 \times 10^{-9},$$

where $p_n^{(N)} \in C[0, 1]$ is the piecewise Lagrange interpolation polynomial of f defined as in the previous problem.

17. Let x_1, \dots, x_n be n distinct points and l_1, \dots, l_n the associated Lagrange basis polynomials. Prove the identity

$$\sum_{j=1}^n (x - x_j)^2 l_j'(x_j) l_j^2(x) = 0.$$

18. Let $N \geq 1$ be an integer and $x_j = j2\pi/N$, $j = 0, \dots, N-1$. Show for any integers k and l that

$$\sum_{j=0}^{N-1} e^{ikx_j} e^{-ilx_j} = \begin{cases} N & \text{if } k \equiv l \pmod{N}, \\ 0 & \text{if } k \not\equiv l \pmod{N}. \end{cases}$$

19. Let $n \geq 0$ be an integer,

$$x_0 = \frac{\pi}{2(n+1)}, \quad \text{and} \quad x_j = x_0 + \frac{j\pi}{2(n+1)}, \quad j = 1, \dots, n.$$

Let $g \in C[0, \pi]$. Prove that there exists a unique $T_n \in \text{span}\{1, \cos x, \dots, \cos nx\}$ such that

$$T_n(x_j) = g(x_j) \quad j = 0, \dots, n.$$

Moreover,

$$T_n(x) = \frac{\gamma_0}{2} + \sum_{k=1}^n \gamma_k \cos kx,$$

where

$$\gamma_k = \frac{2}{n+1} \sum_{j=0}^n g(x_j) \cos kx_j, \quad k = 0, \dots, n.$$

20. Let $N \geq 1$ be an integer. Let

$$\Pi_N = \left\{ (a_k)_{k=-\infty}^{\infty} : a_k \in \mathbb{C}, a_{k+N} = a_k, \forall k = 0, \pm 1, \dots \right\}$$

denote the space of all bi-infinite, N -periodic complex sequences. For any $\mathbf{a} = (a_k)_{k=-\infty}^{\infty} \in \Pi_N$ and $\mathbf{b} = (b_k)_{k=-\infty}^{\infty} \in \Pi_N$, define the convolution $\mathbf{c} = \mathbf{a} * \mathbf{b} \in \Pi_N$ by $\mathbf{c} = (c_k)_{k=-\infty}^{\infty}$ with

$$c_k = \frac{1}{N} \sum_{j=0}^{N-1} a_j b_{k-j} \quad \forall k = 0, \pm 1, \dots$$

Prove that the discrete Fourier transform converts convolution into multiplication:

$$(\mathcal{F}_N \mathbf{c})_k = (\mathcal{F}_N \mathbf{a})_k (\mathcal{F}_N \mathbf{b})_k \quad \forall k = 0, \pm 1, \dots$$

21. Let $n \geq 1$ be an integer and $\Delta = \{a = x_0 < \dots < x_n = b\}$. Suppose that $s \in \mathcal{S}_3(\Delta)$ and $f \in H^4(a, b)$ satisfy

$$s(x_j) = f(x_j) \quad j = 0, \dots, n.$$

Suppose also that one of the following conditions is satisfied:

- (a) $s'(a) = f'(a)$ and $s'(b) = f'(b)$;
- (b) $s''(a) = f''(a)$ and $s''(b) = f''(b)$;
- (c) $f \in H_p^4(a, b)$ and $s \in H_p^3(a, b)$.

Show that

$$\int_a^b [f''(x) - s''(x)]^2 dx = \int_a^b [f(x) - s(x)] f^{(4)}(x) dx.$$

22. Let $\Delta = \{a = x_0 < \dots < x_n = b\}$ be a partition of $[a, b]$. Consider the boundary condition

$$s^{(k)}(x_0) = s^{(k)}(x_n) = 0, \quad k = 0, 1, 2.$$

- (a) Show that any cubic spline on Δ satisfying the given boundary condition vanishes identically if $1 \leq n \leq 3$.
 - (b) Show that any cubic spline on Δ satisfying the given boundary condition is uniquely determined by its value at x_2 if $n = 4$.
 - (c) Let $n = 4$ and $x_j = -2, -1, 0, 1, 2$. Find explicitly the cubic spline $s \in \mathcal{S}_3(\Delta)$ that satisfies that given boundary condition and that $s(0) = 1$.
23. Let $n \geq 1$ be an integer and $\Delta = \{a = x_0 < \dots < x_n = b\}$ be a partition of $[a, b]$. Denote by \mathcal{S} the set of all cubic splines s on Δ that satisfy $s''(x_0) = s''(x_n) = 0$.

- (a) Show that for each j with $0 \leq j \leq n$, there exists a unique $S_j \in \mathcal{S}$ that satisfies

$$S_j(x_k) = \delta_{jk} \quad k = 0, \dots, n.$$

- (b) Let $f \in C[a, b]$. Show that

$$S(x) = \sum_{j=0}^n f(x_j) S_j(x)$$

is the unique spline in \mathcal{S} that interpolates f at x_0, \dots, x_n .

- (c) What is the dimension of \mathfrak{S} ?
24. Let $x_j = a + jh$ ($j = 0, \dots, n$) with $h > 0$. Denote by $\mathfrak{S}_3(\Delta)$ the set of splines determined by these knots x_j ($j = 0, \dots, n$). Define $S_j \in \mathfrak{S}_3(\Delta)$ ($j = 0, \dots, n$) by

$$S_j(x_k) = \delta_{jk} \quad j, k = 0, \dots, n \quad \text{and} \quad S_j''(x_0) = S_j''(x_n) = 0.$$

Fix j with $0 \leq j \leq n$. Show that the moments M_1, \dots, M_{n-1} of S_j are given by

$$M_i = \begin{cases} -\frac{1}{\rho_i} M_{i+1} & i = 1, \dots, j-2, \\ -\frac{1}{\rho_{n-i}} M_{i-1} & i = j+2, \dots, n-1, \end{cases}$$

$$\begin{cases} M_j = -\frac{6(2 + 1/\rho_{j-1} + 1/\rho_{n-j-1})}{h^2(4 - 1/\rho_{j-1} - 1/\rho_{n-j-1})} \\ M_{j-1} = \frac{6h^{-2} - M_j}{\rho_{j-1}} \\ M_{j+1} = \frac{6h^{-2} - M_j}{\rho_{n-j-1}} \end{cases} \quad j \neq 0, 1, n-1, n,$$

where

$$\rho_1 = 4 \quad \text{and} \quad \rho_i = 4 - 1/\rho_{i-1} \quad i = 2, \dots$$

Chapter 3

Numerical Integration

3.1 The Basics

Definition 3.1 (Numerical quadrature). Let x_1, \dots, x_n be n distinct points in $[a, b]$ and $A_1, \dots, A_n \in \mathbb{R}$. We call

$$\int_a^b f(x) dx \approx \sum_{k=1}^n A_k f(x_k) \quad (3.1)$$

a numerical quadrature with x_1, \dots, x_n quadrature points and A_1, \dots, A_n coefficients.

We say that the quadrature (3.1) is exact for an integrable function $f : [a, b] \rightarrow \mathbb{R}$, if

$$\int_a^b f(x) dx = \sum_{k=1}^n A_k f(x_k).$$

Definition 3.2 (Degree of precision). The degree of precision of a numerical quadrature (3.1) is the smallest integer $m \geq 0$ such that the quadrature (3.1) is exact for $f_j(x) = x^j$, $j = 0, \dots, m$, but is not exact for $f_{m+1}(x) = x^{m+1}$.

Equivalently, the degree of precision of (3.1) is m if and only if (3.1) is exact for all $f \in \mathcal{P}_m$ but is not exact for some $f \in \mathcal{P}_{m+1}$.

Since the quadrature (3.1) is determined by $2n$ parameters x_1, \dots, x_n and A_1, \dots, A_n , we expect that the degree of precision of (3.1) can not exceed $2n - 1$. This is in fact the case.

Proposition 3.3. The degree of precision of any numerical quadrature (3.1) is $\leq 2n - 1$.

Proof. Let $p_{2n}(x) = \prod_{k=1}^n (x - x_k)^2$. Then $p_{2n} \in \mathcal{P}_{2n}$. Moreover, $\int_a^b p_{2n}(x) dx > 0$ and $\sum_{k=1}^n A_k p_{2n}(x_k) = 0$. Thus, the degree of precision $\leq 2n - 1$. \square

If the quadrature points of the numerical quadrature (3.1) are known, then one can use the *method of determined coefficients* to find the coefficients A_1, \dots, A_n so that the degree of precision of the quadrature can be as high as possible.

An example of the method of undetermined coefficients. Find A_1 and A_2 such that the numerical quadrature

$$\int_0^1 f(x)dx \approx A_1f(0) + A_2f(1)$$

has the degree of precision as high as possible.

We choose A_1 and A_2 so that this quadrature is exact for $f(x) = 1$ and $f(x) = x$:

$$\begin{aligned} \int_0^1 dx &= 1 = A_1 + A_2, \\ \int_0^1 x dx &= \frac{1}{2} = A_2. \end{aligned}$$

Solving these two equations, we obtain that $A_1 = A_2 = 1/2$. The quadrature thus becomes

$$\int_0^1 f(x)dx \approx \frac{1}{2}[f(0) + f(1)].$$

To find out the degree of precision of this quadrature, we check its exactness for $f(x) = x^2$. We have

$$\begin{aligned} \int_0^1 x^2 dx &= \frac{1}{3}, \\ \frac{1}{2}(0^2 + 1^2) &= \frac{1}{2}. \end{aligned}$$

Thus, this quadrature is not exact for $f(x) = x^2$. Consequently, the degree of precision of this quadrature is 1.

In the rest of this section, we give a few examples of simple numerical quadrature. For each of these examples, we determine the degree of precision, give the related composite formula, and derive its error formula.

The left-endpoint rectangle rule.

$$\int_a^b f(x)dx \approx f(a)(b-a). \tag{3.2}$$

$$\int_a^b 1dx = 1(b-a)$$

$$\int_a^b x dx = \frac{1}{2}(b^2 - a^2) \neq a(b - a)$$

So, the degree of precision is $m = 0$.

$$\begin{aligned} f \in C[a, b] : \int_a^b f(x) dx - f(a)(b - a) &= \int_a^b [f(x) - f(a)] dx \\ &= \int_a^b f'(\xi(x))(x - a) dx = f'(\xi) \int_a^b (x - a) dx \\ &= \sum_{j=1}^N \frac{1}{2} h \cdot \frac{b - a}{N} \cdot f'(\xi_j) \\ &= \frac{1}{2}(b - a) f'(\xi), \quad \xi \in [a, b] \end{aligned}$$

1

Composite left-endpoint rectangular rule

Let $f \in C[a, b]$, $N \geq 1$ and integer, and $h = \frac{b-a}{N}$. Let

$$x_j = a + jh, \quad j = 0, \dots, N$$

$$\begin{aligned} \int_a^b f(x) dx &= \sum_{j=1}^N \int_{x_{j-1}}^{x_j} f(x) dx \\ &\approx \sum_{j=1}^N f(x_{j-1}) \cdot (x_j - x_{j-1}) \\ &= h \sum_{j=1}^N f(x_{j-1}) = h \sum_{j=0}^{N-1} f(x_j) \end{aligned}$$

Error:

$$f \in C^1[a, b]$$

¹**Generalized Mean-Value Theorem for Integrals.** Let $u \in C[a, b]$. Let $v : [a, b] \rightarrow \mathbb{R}$ be integrable with $v(x) \geq 0$ for all $x \in [a, b]$ or $v(x) \leq 0$ for all $x \in [a, b]$. Then there exists $\xi \in [a, b]$ such that

$$\int_a^b u(x)v(x) dx = u(\xi) \int_a^b v(x) dx.$$

$$\begin{aligned}
& \int_a^b f(x)dx - h \sum_{j=0}^{N-1} f(x_j) \\
&= \sum_{j=1}^N \left[\int_{x_{j-1}}^{x_j} f(x)dx - f(x_{j-1})h \right] \\
&= \sum_{j=1}^N \frac{1}{2}h^2 f'(\xi_j) \\
&= \sum_{j=1}^N \frac{1}{2}h \cdot \frac{b-a}{N} \cdot f'(\xi_j) \\
&= \frac{1}{2}(b-a)hf'(\xi).
\end{aligned}$$

Since

$$\min_{x \in [a,b]} f'(x) \leq \frac{1}{N} \sum_{j=1}^N f'(\xi_j) \leq \max_{x \in [a,b]} f'(x),$$

it follows from the Intermediate-Value Theorem ² that

The midpoint rectangular rule.

$$\int_a^b f(x)dx \approx f\left(\frac{a+b}{2}\right)(b-a)$$

$$\begin{aligned}
\int_a^b 1dx &= 1(b-a) \\
\int_a^b xdx &= \frac{1}{2}(b^2 - a^2) = \left(\frac{a+b}{2}\right)(b-a) = \frac{1}{2}(b^2 - a^2) \\
\int_a^b x^2dx &= \frac{1}{3}(b^3 - a^3) \neq \left(\frac{a+b}{2}\right)^2(b-a)
\end{aligned}$$

$$\begin{aligned}
& \frac{1}{3}(b-a)(b^2 - ba + a^2) \neq \frac{1}{4}(a+b)^2(b-a) \\
& \Leftrightarrow \frac{1}{3}(b^2 - ba + a^2) \neq \frac{1}{4}(a^2 + 2ab + b^2) \\
& \Leftrightarrow 4b^2 - 4ba + 4a^2 \neq 3a^2 + 6ab + 3b^2
\end{aligned}$$

²**The Intermediate-Value Theorem.** If $f \in C[a, b]$ and $\mu \in \mathbb{R}$ satisfy $\min_{x \in [a,b]} f(x) \leq \mu \leq \max_{x \in [a,b]} f(x)$, then there exist $\xi \in [a, b]$ such that $f(\xi) = \mu$.

$$\Leftrightarrow b^2 + a^2 - 2ab \neq 0 \Leftrightarrow a \neq b$$

So the degree of precision is $m = 1$.

Let $f \in C^2[a, b]$

$$\begin{aligned} & \int_a^b f(x)dx - f\left(\frac{a+b}{2}\right)(b-a) \\ &= \int_a^b \left[f(x) - f\left(\frac{a+b}{2}\right)\right]dx \\ &= \int_a^b \left[f\left(\frac{a+b}{2}\right)\left(x - \frac{a+b}{2}\right) + \frac{1}{2}f''(\xi(x))\left(x - \frac{a+b}{2}\right)^2\right]dx \\ &= \frac{1}{2} \int_a^b f''(\xi(x))\left(x - \frac{a+b}{2}\right)^2 dx \\ &= \frac{1}{2} f''(\xi) \int_a^b \left(x - \frac{a+b}{2}\right)^2 dx \\ &= \frac{1}{24} (b-a)^3 f''(\xi). \end{aligned}$$

The composite mid-point rule

$$\begin{aligned} \int_a^b f(x)dx &= \sum_{j=1}^N \int_{x_{j-1}}^{x_j} f(x)dx \\ &\approx \sum_{j=1}^N f\left(\frac{x_{j-1} + x_j}{2}\right)(x_j - x_{j-1}) \\ &= h \sum_{j=1}^N f(x_{j-\frac{1}{2}}) \quad x_{j-\frac{1}{2}} = \frac{1}{2}(x_{j-1} + x_j) \end{aligned}$$

Error: $f \in C^2[a, b]$

$$\begin{aligned} & \int_a^b f(x)dx - h \sum_{j=1}^N f(x_{j-\frac{1}{2}}) \\ &= \sum_{j=1}^N \left[\int_{x_{j-1}}^{x_j} f(x)dx - f\left(\frac{x_{j-1} + x_j}{2}\right)h \right] \\ &= \sum_{j=1}^N \frac{1}{24} h^3 f''(\xi_j) \\ &= \frac{(b-a)}{24} h^2 f''(\xi) \quad \xi \in [a, b]. \end{aligned}$$

The trapezoidal rule.

$$\begin{aligned}\int_a^b f(x)dx &\approx \frac{1}{2}[f(a) + f(b)](b - a) \\ \int_a^b 1dx &= \frac{1}{2}(1 + 1)(b - a) \\ \int_a^b xdx &= \frac{1}{2}(b^2 - a^2) = \frac{1}{2}[a + b](b - a) \\ \int_a^b x^2dx &= \frac{1}{3}(b^3 - a^3) \neq \frac{1}{2}(a^2 + b^2)(b - a) \\ m &= 1\end{aligned}$$

Let $f \in C^2[a, b]$.

$$\begin{aligned}\int_a^b f(x)dx - \frac{1}{2}[f(a) + f(b)](b - a) \\ = \int_a^b \left\{ f(x) - \frac{1}{2}[f(a) + f(b)] \right\} dx \\ = -\frac{1}{12}(b - a)^3 f''(\xi), \text{ for some } \xi \in [a, b].\end{aligned}$$

Composite Rule

$$\begin{aligned}\int_a^b f(x)dx &\approx \frac{h}{2}[f(x_0) + f(x_N)] + h \sum_{j=1}^{N-1} f(x_j) \\ \int_a^b f(x)dx - \left\{ \frac{h}{2}[f(x_0) + f(x_N)] + h \sum_{j=1}^{N-1} f(x_j) \right\} &= -\frac{(b - a)h^2}{12} f''(\xi)\end{aligned}$$

3.2 Interpolatory Quadrature

Let x_0, \dots, x_n be $n + 1$ distinct points in $[a, b]$. Let $f \in C[a, b]$. The Lagrange interpolation polynomial $L_n f \in \mathcal{P}_n$ of f at x_0, \dots, x_n is given by

$$(L_n f)(x) = \sum_{k=0}^n f(x_k) l_k(x),$$

where $l_k(x)$ ($k = 0, \dots, n$) are the Lagrange basis polynomials associated with x_0, \dots, x_n .

$$l_k(x) = \prod_{\substack{i=0 \\ i \neq k}}^n \frac{x - x_i}{x_k - x_i}, \quad k = 0, \dots, n.$$

The approximation

$$\int_a^b f(x)dx \approx \int_a^b (L_n f)(x)dx = \sum_{k=0}^n \left[\int_a^b l_k(x)dx \right] f(x_k)$$

leads to the following:

Definition 3.4 (Interpolatory quadrature). *The interpolatory quadrature associated with $n + 1$ distinct points x_0, \dots, x_n in $[a, b]$ is the numerical quadrature*

$$\int_a^b f(x)dx \approx \sum_{k=0}^n A_k f(x_k)$$

with

$$A_k = \int_a^b l_k(x)dx, \quad k = 0, \dots, n, \quad (3.3)$$

where l_0, \dots, l_n are the Lagrange basis polynomials associated with x_0, \dots, x_n .

Theorem 3.5 (Characterization of interpolatory quadrature). *Let x_0, \dots, x_n be $n+1$ distinct points in $[a, b]$. A numerical quadrature*

$$\int_a^b f(x)dx \approx \sum_{k=0}^n B_k f(x_k) \quad (3.4)$$

is an interpolatory quadrature if and only if its degree of precision is $\geq n$.

Proof. The “if” part. Suppose the degree of precision of (3.4) is $\geq n$. Let $l_0, \dots, l_n \in \mathcal{P}_n$ be the Lagrange basis polynomials associated with x_0, \dots, x_n . Since each $l_j \in \mathcal{P}_n$, we have

$$\int_a^b l_j(x)dx = \sum_{k=0}^n B_k l_j(x_k) = \sum_{k=0}^n B_k \delta_{jk} = B_j.$$

Thus, the quadrature (3.4) is interpolatory.

The “only if” part. Suppose (3.4) is interpolatory. Then the coefficients are given by

$$B_k = \int_a^b l_k(x)dx, \quad k = 0, \dots, n.$$

Let $f \in \mathcal{P}_n$. Then $L_n f = f$ by Proposition 2.5. Consequently,

$$\int_a^b f(x)dx = \int_a^b (L_n f)(x)dx = \sum_{k=0}^n \int_a^b l_k(x)dx f(x_k) = \sum_{k=0}^n B_k f(x_k).$$

This implies that the degree of precision of (3.4) is $\geq n$. □

Definition 3.6 (Newton-Cotes formula). Let $n \geq 1$ be an integer. A (closed) Newton-Cotes formula is an interpolatory quadrature

$$\int_a^b f(x)dx \approx \sum_{k=0}^n A_k f(x_k) \quad (3.5)$$

with the quadrature points $x_k = a + k(b - a)/n$ ($k = 0, \dots, n$).

By the definition of interpolatory quadrature, the coefficients A_k ($k = 0, \dots, n$) in the Newton-Cotes formula (3.5) are given by (3.3) with $l_k \in \mathcal{P}_n$ ($k = 0, \dots, n$) the Lagrange basis polynomials associated with the evenly distributed quadrature points x_k ($k = 0, \dots, n$).

Examples. (1) Consider the Newton-Cotes formula with $n = 1$. We have $x_0 = a$, $x_1 = b$, and

$$A_0 = \int_a^b l_0(x)dx = \int_a^b \frac{x-b}{a-b}dx = \frac{1}{2}(b-a),$$

$$A_1 = \int_a^b l_1(x)dx = \int_a^b \frac{x-a}{b-a}dx = \frac{1}{2}(b-a).$$

Thus, the formula is

$$\int_a^b f(x)dx \approx \frac{1}{2}(b-a)f(x_0) + \frac{1}{2}(b-a)f(x_1) = \frac{1}{2}(b-a)[f(x_0) + f(x_1)].$$

This is exactly the trapezoidal rule.

(2) Consider the Newton-Cotes formula with $n = 2$. We have $x_0 = a$, $x_1 = (a+b)/2 =: c$, $x_2 = b$, and

$$A_0 = \int_a^b l_0(x)dx = \int_a^b \frac{(x-c)(x-b)}{(a-c)(a-b)}dx = \frac{b-a}{6},$$

$$A_1 = \int_a^b l_1(x)dx = \int_a^b \frac{(x-a)(x-b)}{(c-a)(c-b)}dx = \frac{2(b-a)}{3},$$

$$A_2 = \int_a^b l_2(x)dx = \int_a^b \frac{(x-a)(x-c)}{(b-a)(b-c)}dx = \frac{b-a}{6}.$$

The formula is

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

This is called *Simpson's rule*.

In the case $[a, b] = [-1, 1]$, this becomes

$$\int_{-1}^1 f(x)dx \approx \frac{1}{3}[f(-1) + 4f(0) + f(1)].$$

We can verify directly that this is exact for $f(x) = 1, x, x^2$. In fact, it is also exact for $f(x) = x^3$ but not for $f(x) = x^4$. Therefore, the degree of precision of Simpson's rule is 3.

Theorem 3.7 (Error formula for Newton-Cotes formula). *Consider a Newton-Cotes formula*

$$\int_a^b f(x)dx \approx \sum_{k=0}^n A_k f(x_k). \quad (3.6)$$

(1) *If n is even and $f \in C^{n+2}[a, b]$, then exists $\xi \in (a, b)$ such that*

$$\int_a^b f(x)dx - \sum_{k=0}^n A_k f(x_k) = \frac{f^{(n+2)}(\xi)}{(n+2)!} \mu_n, \quad (3.7)$$

where

$$\mu_n = \int_a^b x(x-x_0) \cdots (x-x_n) dx < 0.$$

(2) *If n is odd and $f \in C^{n+1}[a, b]$, then there exists $\eta \in (a, b)$ such that*

$$\int_a^b f(x)dx - \sum_{k=0}^n A_k f(x_k) = \frac{f^{(n+1)}(\eta)}{(n+1)!} \nu_n, \quad (3.8)$$

where

$$\nu_n = \int_a^b (x-x_0) \cdots (x-x_n) dx < 0.$$

Colloary 3.8. *The degree of precision of the Newton-Cotes formula (3.6) with quadrature points $x_k = a + k(b-a)/n$ ($k = 0, \dots, n$) is n if n is odd and $n+1$ if n is even.*

Proof. Suppose $n \geq 1$ is even. By (3.7), the quadrature (3.6) is exact for all $f \in \mathcal{P}_{n+1}$. Setting $f(x) = x^{n+2}$ in (3.7), we see that the right-hand side of (3.7) is $\mu_n \neq 0$. Thus, the degree of precision in this case is $n+1$. The same argument applies to the case n is odd. \square

To prove Theorem 3.7, we first prove the following:

Lemma 3.9. *Let $n \geq 1$ be an even number, $h = (b-a)/n$, and $x_k = a + kh$ ($k = 0, \dots, n$).*

Let

$$\omega_n(x) = (x-x_0) \cdots (x-x_n) \quad \text{and} \quad \Omega_n(x) = \int_a^x \omega_n(t)dt. \quad (3.9)$$

Then $\Omega_n(a) = \Omega_n(b) = 0$ and $\Omega_n(x) > 0$ for all $x \in (a, b)$.

Proof. It is obvious that $\Omega_n(a) = \int_a^a \omega_n(t)dt = 0$. Since n is even, we have by the change of variable $x = a + h(t + n/2)$ that

$$\Omega_n(b) = \int_a^b \omega(x)dx = h^{n+2} \int_{-\frac{n}{2}}^{\frac{n}{2}} t \prod_{k=1}^{\frac{n}{2}} (t^2 - k^2)dt = 0.$$

Let an index j be such that $1 \leq j \leq n/2$. We claim:

$$|\omega_n(x)| > |\omega_n(x+h)| \quad \forall x \in (x_{2j-2}, x_{2j-1}). \quad (3.10)$$

To see this, let us fix an $x \in (x_{2j-2}, x_{2j-1})$. Let $x = a + ht$ for some $t \in \mathbb{R}$. Clearly, $2j-2 < t < 2j-1$. Thus, t is not an integer. Moreover, since $1 \leq j \leq n/2$, we have $0 < t+1 < n/2$. From $x = a + ht$ and $x_k = a + kh$ ($k = 0, \dots, n$), we then obtain that

$$\left| \frac{\omega_n(x+h)}{\omega_n(x)} \right| = \left| \frac{(t+1)t \cdot (t-1) \dots (t-n+1)}{t(t-1) \dots (t-n)} \right| = \left| \frac{t+1}{t-n} \right| = \frac{t+1}{n-t} < 1,$$

where the last inequality is equivalent to the true fact that $t+1/2 < n/2$. This proves (3.10).

Consider now $[x_0, x_2] = [x_0, x_1] \cup [x_1, x_2]$. Since n is even, $\omega_n(x) < 0$ on $(-\infty, x_0)$ and $\omega_n(x) > 0$ on (x_0, x_1) . Thus, by the fact that $\Omega_n(x_0) = 0$, we have $\Omega_n(x) > 0$ on $(x_0, x_1]$. Let $x \in (x_1, x_2]$. Then $x-h \in (x_0, x_1]$. Hence, $\omega_n(t) > 0$ for any $t \in (x_0, x-h)$. This and (3.10) with $j=1$ imply that $\omega_n(t) + \omega_n(t+h) > 0$ for any $t \in (x_0, x-h)$. Therefore, by the change of variable $s = t-h$ and (3.10) for $j=1$,

$$\begin{aligned} \Omega_n(x) &= \int_a^x \omega_n(t) dt = \int_{x_0}^{x_1} \omega_n(t) dt + \int_{x_1}^x \omega_n(t) dt \\ &\geq \int_{x_0}^{x-h} \omega_n(t) dt + \int_{x_0}^{x-h} \omega_n(s+h) ds = \int_{x_0}^{x-h} [\omega_n(t) + \omega_n(t+h)] dt > 0. \end{aligned}$$

Hence, $\Omega_n(x) > 0$ on $[x_0, x_2]$. Since $\Omega'_n(x) = \omega_n(x) > 0$ in (x_2, x_3) and $\Omega_n(x_2) > 0$, we have $\Omega_n(x) > 0$ on $[x_2, x_3]$. A similar argument then leads to $\Omega_n(x) > 0$ on $[x_2, x_4]$. Continuing this process, we have $\Omega_n(x) > 0$ for $x \in (x_0, x_{2j-1}]$ with $2j-1 = n/2$ or $2j-1 = (n/2)-1$. In the latter case, we have $\Omega_n(x) > 0$ in $(x_{2j-1}, x_{n/2})$, since $\Omega'_n(x) = \omega_n(x) > 0$ in $(x_{2j-1}, x_{n/2})$ and $\Omega_n(x) > 0$ on $(x_0, x_{2j-1}]$. Therefore, $\Omega_n(x) > 0$ on $(x_0, x_{n/2}]$.

If $x \in (x_{n/2}, x_n)$ then $x_{n/2} - (x - x_{n/2}) \in (x_0, x_{n/2})$ and hence $\Omega_n(x_{n/2} - (x - x_{n/2})) > 0$. Moreover, by the change of variable $s = t - x_{n/2}$, we have

$$\int_{x_{n/2} - (x - x_{n/2})}^{x_{n/2} + (x - x_{n/2})} \omega_n(t) dt = \int_{-(x - x_{n/2})}^{x - x_{n/2}} \omega_n(s + x_{n/2}) ds = \int_{-(x - x_{n/2})}^{x - x_{n/2}} s \prod_{k=1}^{n/2} [s^2 - (kh)^2] ds = 0.$$

Therefore,

$$\begin{aligned} \Omega_n(x) &= \int_{x_0}^x \omega_n(t) dt \\ &= \int_{x_0}^{x_{n/2} - (x - x_{n/2})} \omega_n(t) dt + \int_{x_{n/2} - (x - x_{n/2})}^{x_{n/2} + (x - x_{n/2})} \omega_n(t) dt \\ &= \Omega_n(x_{n/2} - (x - x_{n/2})) > 0. \end{aligned}$$

The proof is complete. \square

Proof of Theorem 3.7. Let ω_n and Ω_n be given as in (3.9). Since (3.6) is an interpolatory formula, we have by Theorem 2.10 and Proposition ? (on the properties of divided differences with repeated points) that

$$\begin{aligned} e_n(f) &:= \int_a^b f(x)dx - \sum_{k=0}^n A_k f(x_k) \\ &= \int_a^b f(x)dx - \int_a^b (L_n f)(x)dx \\ &= \int_a^b f[x_0, \dots, x_n, x] \omega_n(x) dx \quad \forall f \in C^1[a, b]. \end{aligned} \quad (3.11)$$

Case 1: n is even and $f \in C^{n+2}[a, b]$. By integration by parts, we obtain from (3.11), Lemma 3.9, Proposition (?) (on properties on divided differences with repeated points), and the Generalized Mean-Value Theorem for integrals that

$$\begin{aligned} e_n(f) &= \int_a^b f[x_0, \dots, x_n, x] \Omega'_n(x) dx \\ &= f[x_0, \dots, x_n, x] \Omega_n(x) \Big|_{x=a}^{x=b} - \int_a^b \Omega(x) \frac{d}{dx} f[x_0, \dots, x_n, x] dx \\ &= - \int_a^b \Omega_n(x) \frac{d}{dx} f[x_0, \dots, x_n, x] dx \\ &= - \int_a^b \Omega_n(x) \frac{f^{(n+2)}(\xi(x)) dx}{(n+2)!} \\ &= - \frac{f^{(n+2)}(\xi)}{(n+2)!} \int_a^b \Omega_n(x) dx \end{aligned}$$

for some $\xi \in (a, b)$. By integration by parts and Lemma 3.9, we then have

$$\int_a^b \Omega_n(x) dx = x \Omega_n(x) \Big|_a^b - \int_a^b x \Omega'_n(x) dx = - \int_a^b x \omega_n(x) dx = -\mu_n.$$

Since $\int_a^b \Omega_n(x) dx > 0$ by Lemma 3.9, $\mu_n < 0$.

Case 2: n is odd and $f \in C^{n+1}[a, b]$. It follows from (3.11) that

$$e_n(f) = \int_a^{b-h} \omega_n(x) f[x_0, \dots, x_n, x] dx + \int_{b-h}^b \omega_n(x) f[x_0, \dots, x_n, x] dx =: I_1 + I_2. \quad (3.12)$$

Since $\omega_n(x)$ does not change sign in $(b-h, b)$, we have by the Generalized Mean-Value Theorem for integrals that

$$I_2 := \int_{b-h}^b \omega_n(x) f[x_0, \dots, x_n, x] dx = \frac{f^{(n+1)}(\eta')}{(n+1)!} \int_{b-h}^b \omega_n(x) dx \quad (3.13)$$

for some $\eta' \in (a, b)$.

By the definition of divided differences, we have

$$\begin{aligned}
I_1 &:= \int_a^{b-h} \omega_n(x) f[x_0, \dots, x_n, x] dx \\
&= \int_a^{b-h} \omega_{n-1}(x)(x - x_n) \left(\frac{f[x_0, \dots, x_{n-1}, x] - f[x_0, \dots, x_{n-1}, x_n]}{x - x_n} \right) dx \\
&= \int_a^{b-h} \Omega'_{n-1}(x) f[x_0, \dots, x_{n-1}, x] dx - \int_a^{b-h} \Omega'_{n-1}(x) f[x_0, \dots, x_{n-1}, x_n] dx \\
&=: J_1 - J_2.
\end{aligned}$$

Since $n - 1$ is even, by Lemma 3.9 we have $\Omega_{n-1}(a) = \Omega_{n-1}(b - h) = 0$. Consequently,

$$J_2 = f[x_0, \dots, x_{n-1}, x_n] \int_a^{b-h} \Omega'_{n-1}(x) dx = f[x_0, \dots, x_{n-1}, x_n] [\Omega_{n-1}(b - h) - \Omega_{n-1}(a)] = 0. \quad (3.14)$$

Again by Lemm 3.9, $\Omega_{n-1}(x) > 0$ on $(a, b - h)$. By integration by parts, Propertion? (on properties of divided differences with repeated points), and the Generalized Mean-Value Theorem for integrals, we get

$$\begin{aligned}
J_1 &:= \int_a^{b-h} \Omega'_{n-1}(x) f[x_0, \dots, x_{n-1}, x] dx \\
&= - \int_a^{b-h} \Omega_{n-1}(x) \frac{d}{dx} f[x_0, \dots, x_{n-1}, x] dx \\
&= - \frac{f^{(n+1)}(\eta'')}{(n+1)!} \int_a^{b-h} \Omega_{n-1}(x) dx \quad (3.15)
\end{aligned}$$

for some $\eta'' \in (a, b)$.

From (3.12)–(3.15), we obtain

$$e_n(f) = -[A f^{(n+1)}(\eta') + B f^{(n+1)}(\eta'')],$$

where

$$\begin{aligned}
A &= -\frac{1}{(n+1)!} \int_{b-h}^b \omega_n(x) dx, \\
B &= \frac{1}{(n+1)!} \int_a^{b-h} \Omega_{n-1}(x) dx.
\end{aligned}$$

Clearly, $\omega_n(x) > 0$ for any $x > b$. Thus, $\omega_n(x) < 0$ in $(b - h, b)$. This implies that $A > 0$. Since $n - 1$ is even, we have $B > 0$ by Lemma 3.9. The fact that

$$\min_{a \leq x \leq b} f^{(n+1)}(x) \leq \frac{A f^{(n+1)}(\eta') + B f^{(n+1)}(\eta'')}{A + B} \leq \max_{a \leq x \leq b} f^{(n+1)}(x)$$

and the Intermediate-Value Theorem now imply that

$$e_n(f) = -(A + B)f^{(n+1)}(\eta) \quad (3.16)$$

for some $\eta \in (a, b)$. Again since $n - 1$ is even, we obtain by Lemma 3.9 that

$$\begin{aligned} \int_a^{b-h} \omega_n(x) dx &= \int_a^{b-h} \Omega'_{n-1}(x)(x-b) dx \\ &= \Omega_{n-1}(x)(x-b) \Big|_a^{b-h} - \int_a^{b-h} \Omega_{n-1}(x) dx \\ &= - \int_a^{b-h} \Omega_{n-1}(x) dx. \end{aligned}$$

Consequently,

$$\begin{aligned} A + B &= \frac{1}{(n+1)!} \int_{b-h}^b \omega_n(x) dx + \frac{1}{(n+1)!} \int_a^{b-h} \Omega_{n-1}(x) dx \\ &= \frac{1}{(n+1)!} \int_{b-h}^b \omega_n(x) dx - \frac{1}{(n+1)!} \int_a^{b-h} \omega_n(x) dx \\ &= - \frac{1}{(n+1)!} \int_a^b \omega_n(x) dx. \end{aligned} \quad (3.17)$$

This, together (3.16), leads to (3.8). Since A and B are positive, then we have by (3.17) that

$$\nu_n = \int_a^b \omega_n(x) dx = -(n+1)!(A+B) < 0,$$

completing the proof. \square

3.3 Peano Kernel and Error Representation

Theorem 3.10 (Peano kernel and error representation for numerical quadrature).

Assume the degree of precision of a given numerical quadrature

$$\int_a^b f(x) dx \approx \sum_{k=0}^n A_k f(x_k)$$

is m . Then

$$\int_a^b f(x) - \sum_{k=0}^n A_k f(x_k) = \int_a^b \tilde{K}_m(t) f^{(m+1)}(t) dt \quad \forall f \in C^{(m+1)}[a, b], \quad (3.18)$$

where

$$\tilde{K}_m(t) = \frac{1}{m!} \left[\int_a^b (x-t)_+^m dx - \sum_{k=0}^n A_k (x_k - t)_+^m \right]. \quad (3.19)$$

Proof. Let $f \in C^{(m+1)}[a, b]$. By the Taylor expansion,

$$\begin{aligned} f(x) &= f(a) + f'(a)(x-a) + \cdots + \frac{f^{(m)}(a)}{m!}(x-a)^m + \frac{1}{m!} \int_a^b (x-t)_+^m f^{(m+1)}(t) dt \\ &= Q_m(x) + R_m(x). \end{aligned}$$

Since the degree of precision of (3.18) is m and

$$Q_m(x) = f(a) + f'(a)(x-a) + \cdots + \frac{f^{(m)}(a)}{m!}(x-a)^m$$

is a polynomial of degree $\leq m$, the quadrature (3.18) is exact for Q_m . Thus,

$$\begin{aligned} &\int_a^b f(x) dx - \sum_{k=0}^n A_k f(x_k) \\ &= \int_a^b Q_m(x) dx - \sum_{k=0}^n A_k Q_m(x_k) + \int_a^b R_m(x) dx - \sum_{k=0}^n A_k R_m(x_k) \\ &= \int_a^b R_m(x) dx - \sum_{k=0}^n A_k R_m(x_k) \\ &= \frac{1}{m!} \int_a^b \left[\int_a^b (x-t)_+^m dx - \sum_{k=0}^n A_k (x_k - t)_+^m \right] f^{(m+1)}(t) dt \\ &= \int_a^b \tilde{K}_m(t) f^{(m+1)}(t) dt. \end{aligned}$$

The proof is complete. □

We call \tilde{K}_m defined in (3.19) the Peano kernel for the numerical quadrature (3.18).

3.4 Euler-Maclaurin Formula and Romberg Algorithm

3.5 Weighted Gaussian Quadrature

Let ρ be a weight function on $[a, b]$. We consider a numerical quadrature

$$\int_a^b \rho(x) f(x) dx \approx \sum_{k=1}^n A_k f(x_k), \quad (3.20)$$

where $n \geq 1$ is an integer, $x_1, \dots, x_n \in [a, b]$ are n distinct points, and $A_1, \dots, A_n \in \mathbb{R}$. It follows from Proposition 3.3 that the degree of precision of any numerical quadrature (3.20)

is less than or equal to $2n - 1$. On the other hand, by ?, the degree of precision of (3.20) is greater than or equal to $n - 1$, if it is interpolatory, i.e., the coefficients are given by

$$A_k = \int_a^b \rho(x) l_k(x) dx, \quad k = 1, \dots, n, \quad (3.21)$$

where

$$l_k(x) = \prod_{\substack{j=1 \\ j \neq k}}^n \frac{x - x_j}{x_k - x_j}, \quad k = 1, \dots, n,$$

are the Lagrange basis polynomials associated with x_1, \dots, x_n .

We want to choose the quadrature points x_1, \dots, x_n so that the formula (3.20) has the degree of precision as high as possible. Let $f \in C[a, b]$. Let $L_{n-1} : C[a, b] \rightarrow \mathcal{P}_{n-1}$ be the Lagrange interpolator associated with x_1, \dots, x_n . By Theorem 2.10, we have the error

$$\begin{aligned} e_n(f) &:= \int_a^b \rho(x) f(x) dx - \sum_{k=1}^n A_k f(x_k) \\ &= \int_a^b \rho(x) [f(x) - (L_{n-1}f)(x)] dx \\ &= \int_a^b \rho(x) f[x_1, \dots, x_n, x] \prod_{k=1}^n (x - x_k) dx. \end{aligned}$$

If $f \in \mathcal{P}_m$ for some integer $m \geq n$, then by Proposition 2.11 the divided difference $f[x_1, \dots, x_n, x]$ is a polynomial in \mathcal{P}_{m-n} . We thus want to choose x_1, \dots, x_n so that

$$\int_a^b \rho(x) p(x) \prod_{k=1}^n (x - x_k) dx = 0 \quad \forall p \in \mathcal{P}_{m-n} \quad (3.22)$$

with possibly $m = n, \dots, 2n - 1$. This will not hold true for $m = 2n$ by Proposition 3.3. It is then clear that if $\omega_n(x) := \prod_{k=1}^n (x - x_k)$ is the n th orthogonal polynomial in $\tilde{\mathcal{P}}_n$, i.e., x_1, \dots, x_n are roots of an n th orthogonal polynomial, then (3.22) will hold true for all $m = n, \dots, 2n - 1$. This means that $e_n(f) = 0$ for any $f \in \mathcal{P}_m$ with $m = n, \dots, 2n - 1$. Clearly, $e_n(f) = 0$ for any $f \in \mathcal{P}_m$ with $m = 0, \dots, n - 1$, if (3.20) is interpolatory. Therefore, the highest possible degree of precision is achieved by an interpolatory quadrature with quadrature points roots of orthogonal polynomials.

Definition 3.11 (Weighted Gaussian quadrature). *A numerical quadrature (3.20) is called a weighted Gaussian quadrature, if*

- (1) *the quadrature points x_1, \dots, x_n are the n simple roots of an n th orthogonal polynomial in $L_\rho^2(a, b)$,*
- (2) *the quadrature is interpolatory, i.e., the coefficients A_1, \dots, A_n are given by (3.21).*

Theorem 3.12 (Characterization of weighted Gaussian quadrature). *A numerical quadrature (3.20) is a weighted Gaussian quadrature if and only if it has the degree of precision $2n - 1$.*

Proof. The “if” part. Assume the degree of precision of the numerical quadrature (3.20) is $2n - 1$. Let $Q_n(x) = \prod_{k=1}^n (x - x_k)$ and $q \in \mathcal{P}_{n-1}$. Then $qQ_n \in \mathcal{P}_{2n-1}$. Since the degree of precision of (3.20) is $2n - 1$,

$$\int_a^b \rho(x)q(x)Q_n(x)dx = \sum_{k=1}^n A_k q(x_k)Q_n(x_k) = 0.$$

Therefore, $Q_n \in \mathcal{P}_n$ is an n th orthogonal polynomial in $L_\rho^2(a, b)$, and hence x_1, \dots, x_n are roots of this polynomial. Moreover, since the quadrature is exact for all polynomials in \mathcal{P}_{n-1} , it is interpolatory by ?. Thus, it is a weighted Gaussian quadrature by Definition 3.11.

The “only if” part. Assume (3.20) is a weighted Gaussian quadrature. For $p \in \mathcal{P}_{2n-1}$, there exist $q \in \mathcal{P}_{n-1}$ and $r \in \mathcal{P}_{n-1}$ with $\deg r < \deg p$ such that

$$p(x) = q(x)\tilde{Q}_n(x) + r(x),$$

where $\tilde{Q}_n(x) = \prod_{k=1}^n (x - x_k)$ in \tilde{P}_n is the n th orthogonal polynomial in $L_\rho^2(a, b)$. Clearly, $p(x_k) = r(x_k)$ for $k = 1, \dots, n$. Thus, by the orthogonality and the fact that the weighted Gaussian quadrature (3.20) is exact for any polynomial in \mathcal{P}_{n-1} , we have

$$\int_a^b \rho(x)p(x) dx = \int_a^b \rho(x)q(x)\tilde{Q}_n(x) dx + \int_a^b \rho(x)r(x) dx = \sum_{k=1}^n A_k r(x_k) = \sum_{k=1}^n A_k p(x_k).$$

Therefore, the degree of precision of (3.20) is greater than or equal to $2n - 1$; and is in fact exactly $2n - 1$ by Proposition 3.3. \square

Theorem 3.13 (Error of weighted Gaussian quadrature). *Let (3.20) be a weighted Gaussian quadrature. For any $f \in C^{2n}[a, b]$, there exists $\xi \in [a, b]$ such that*

$$\int_a^b \rho(x)f(x)dx - \sum_{k=1}^n A_k f(x_k) = \frac{f^{(2n)}(\xi)}{(2n)!} \int_a^b \rho(x) \prod_{k=1}^n (x - x_k)^2 dx.$$

Proof. Let $p \in \mathcal{P}_{2n-1}$ be the Hermite interpolation polynomial of f at x_1, \dots, x_n . Then it follows from Theorem 2.18 that for each $x \in [a, b]$ there exists $\xi(x) \in [a, b]$ such that

$$f(x) - p(x) = \frac{f^{(2n)}(\xi(x))}{(2n)!} \prod_{k=1}^n (x - x_k)^2.$$

Since the weighted Gaussian quadrature (3.20) has the degree of precision $2n - 1$ and since $p(x_k) = f(x_k)$ ($k = 1, \dots, n$), we have by the Generalized Mean-value Theorem for integrals ? that

$$\int_a^b \rho(x)f(x)dx = \int_a^b \rho(x)p(x)dx + \int_a^b \rho(x) \frac{f^{(2n)}(\xi(x))}{(2n)!} \prod_{k=1}^n (x - x_k)^2 dx$$

$$\begin{aligned}
&= \sum_{k=1}^n A_k p(x_k) + \frac{f^{(2n)}(\xi)}{(2n)!} \int_a^b \rho(x) \prod_{k=1}^n (x - x_k)^2 dx \\
&= \sum_{k=1}^n A_k f(x_k) + \frac{f^{(2n)}(\xi)}{(2n)!} \int_a^b \rho(x) \prod_{k=1}^n (x - x_k)^2 dx,
\end{aligned}$$

completing the proof. \square

The following is a useful property of a weighted Gaussian quadrature:

Proposition 3.14. *The coefficients of a weighted Gaussian quadrature are all positive.*

Proof. Let (3.20) be a weighted Gaussian quadrature. Let $l_j \in \mathcal{P}_{n-1}$ ($j = 1, \dots, n$) be the Lagrange basis polynomials associated with the quadrature points $x_1, \dots, x_n \in (a, b)$. Since each $l_j^2 \in \mathcal{P}_{2n-2}$ and the weighted Gaussian quadrature (3.20) has the degree of precision $2n - 1$, we have

$$0 < \int_a^b \rho(x) [l_j(x)]^2 dx = \sum_{k=1}^n A_k [l_j(x_k)]^2 = A_j,$$

completing the proof. \square

Gaussian quadrature

$$\int_{-1}^1 f(x) dx \approx \sum_{k=1}^n A_k f(x_k). \quad (3.23)$$

The Legendre polynomials P_n ($n = 0, \dots$) are orthogonal polynomials in $L^2[-1, 1]$. Recall:

$$\begin{aligned}
P_n(x) &= \frac{1}{2^n n!} \frac{d^n}{dx^n} [(x^2 - 1)^n] \\
\int_{-1}^1 P_m(x) P_n(x) dx &= \begin{cases} 0 & m \neq n \\ \frac{2}{2n+1} & m = n. \end{cases}
\end{aligned}$$

Each P_n ($n \geq 1$) has n roots in $(-1, 1)$.

$$\begin{aligned}
P_0(x) &= 1 \\
P_1(x) &= x \\
P_2(x) &= \frac{1}{2}(3x^2 - 1) \\
P_3(x) &= \frac{1}{2}(5x^3 - 3x) \\
\int_{-1}^1 f(x) dx &\approx 2f(0) \\
\int_{-1}^1 f(x) dx &\approx f\left(-\frac{1}{\sqrt{3}}\right) + f\left(\frac{1}{\sqrt{3}}\right)
\end{aligned}$$

$$\int_{-1}^1 f(x)dx \approx \frac{5}{9}f\left(-\sqrt{\frac{3}{5}}\right) + \frac{8}{9}f(0) + \frac{5}{9}f\left(\sqrt{\frac{3}{5}}\right)$$

Theorem 3.15. *The coefficients of the Gaussian quadrature (3.23) are given by*

$$A_k = \frac{2}{(1-x_k^2)[P'_n(x_k)]^2}, \quad k = 1, \dots, n.$$

Proof. The Lagrange basis polynomials $l_k \in \mathcal{P}_{n-1}$ ($k = 1, \dots, n$) associated with the roots x_1, \dots, x_n of the n th Legendre polynomial P_n are given by

$$l_k(x) = \frac{P_n(x)}{(x-x_k)P'_n(x_k)} = \prod_{\substack{j=1 \\ j \neq k}}^n \frac{x-x_j}{x_k-x_j}, \quad k = 1, \dots, n. \quad (3.24)$$

Fix an integer k with $1 \leq k \leq n$. By integration by parts and the orthogonality, we obtain that

$$S_k := \int_{-1}^1 l_k(x)P'_n(x)dx = l_k(x)P_n(x) \Big|_{-1}^1 - \int_{-1}^1 l'_k(x)P_n(x)dx = l_k(x)P_n(x) \Big|_{-1}^1.$$

Since $l_k P'_n \in \mathcal{P}_{2n-2}$ and the Gaussian quadrature (3.23) has the degree of precision $2n-1$, we also have

$$S_k = \int_{-1}^1 l_k(x)P'_n(x)dx = \sum_{j=1}^n A_j l_k(x_j)P'_n(x_j) = A_k P'_n(x_k).$$

Consequently, by these two equations for S_k , (3.24), and Part (5) of Theorem 1.30, we obtain

$$\begin{aligned} A_k &= \frac{l_k(x)P_n(x) \Big|_{x=-1}^{x=1}}{P'_n(x_k)} = \frac{[P_n(x)]^2 \Big|_{x=-1}^{x=1}}{(x-x_k)[P'_n(x_k)]^2 \Big|_{x=-1}^{x=1}} \\ &= \left(\frac{[P_n(1)]^2}{1-x_k} - \frac{[P_n(-1)]^2}{-1-x_k} \right) \frac{1}{[P'_n(x_k)]^2} = \frac{2}{(1-x_k^2)[P'_n(x_k)]^2}, \end{aligned}$$

completing the proof. □

Remainder

$$\frac{f^{(2n)}(\xi)}{(2n)!} \int_{-1}^1 \prod_{k=1}^n (x-x_k)^2 dx = \frac{f^{(2n)}(\xi)}{(2n)!} \int_{-1}^1 \left(\frac{2^n n!}{(2n)!} \right)^2 P_n^2(x) dx = f^{(2n)}(\xi) \frac{2^{2n+1} (n!)^4}{(2n+1)[(2n)!]^3}$$

The Gauss-Chebyshev quadrature

$$\int_{-1}^1 \frac{f(x)}{\sqrt{1-x^2}} dx \approx \sum_{k=1}^n A_k f(x_k)$$

$x_1, \dots, x_n \in (-1, 1)$ are roots of $T_n(x) = \cos(n \arccos x)$

$$x_k = \cos Q_k = \cos \frac{(2k-1)\pi}{2n}, \quad k = 1, \dots, n$$

$$A_k = \int_{-1}^1 \frac{l_k(x)}{\sqrt{1-x^2}} dx = \int_{-1}^1 \frac{1}{\sqrt{1-x^2}} \left[\frac{T_k(x)}{(x-x_k)T'_k(x_k)} \right] dx$$

$$= \frac{\pi}{n}$$

Proof. We show that the formula is exact for all T_0, \dots, T_{2n-1} . Then, the degree of precision is $\geq 2n-1$. But any numerical quadrature with n quadrature points has degree of precision $\leq 2n-1$. Thus, this has the degree of precision exactly $2n-1$. Therefore, this is the weighted Gaussian quadrature.

Notice that By the change of variable $x = \cos \theta$,

$$\int_{-1}^1 \frac{T_m(x)}{\sqrt{1-x^2}} dx = \int_0^\pi \cos m\theta d\theta = \begin{cases} \pi & \text{if } m = 0, \\ 0 & \text{if } m \neq 0. \end{cases}$$

For $m = 0$, we have

$$\frac{\pi}{n} \sum_{k=1}^n T_0(x_k) = \pi.$$

Thus the formula is exact for $m = 0$.

Consider now $1 \leq m \leq 2n-1$. Denote by i the complex unit, i.e., $i^2 = -1$. We have $e^{im\pi} \neq 1$. Denote by $\mathcal{R}(z)$ the real part of a complex number z . We have

$$\begin{aligned} \sum_{k=1}^n T_m(x_k) &= \sum_{k=1}^n \cos \left(\frac{m(2k-1)\pi}{2n} \right) \\ &= \sum_{k=1}^n \mathcal{R} \left(e^{im(2k-1)\pi/(2n)} \right) \\ &= \mathcal{R} \left(e^{-im\pi/(2n)} \sum_{k=1}^n e^{imk\pi/n} \right) \\ &= \mathcal{R} \left(e^{-im\pi/(2n)} e^{im\pi/n} \frac{e^{im\pi} - 1}{e^{im\pi/n} - 1} \right) \\ &= [(-1)^m - 1] \mathcal{R} \left(\frac{e^{im\pi/(2n)}}{e^{im\pi/n} - 1} \right) \\ &= [(-1)^m - 1] \mathcal{R} \left(\frac{e^{im\pi/(2n)} (e^{-im\pi/n} - 1)}{|e^{im\pi/n} - 1|^2} \right) \\ &= [(-1)^m - 1] \mathcal{R} \left(\frac{e^{-im\pi/(2n)} - e^{im\pi/(2n)}}{|e^{im\pi/n} - 1|^2} \right) \end{aligned}$$

$$= 0.$$

$$\int_{-1}^1 \frac{T_m(x)}{\sqrt{1-x^2}} dx = \frac{\pi}{n} \sum_{k=1}^n T_m(x_k), \quad m = 0, \dots, 2n-1.$$

□

$$\text{Error: } \frac{\pi}{(2n)!2^{2n-1}} f^{(2n)}(\xi).$$

3.6 Convergence of Sequences of Numerical Quadrature

Theorem 3.16 (Convergence of sequences of numerical quadrature). *Given a sequence of numerical quadrature*

$$\int_a^b \rho(x)f(x) dx \approx \sum_{k=1}^n A_k^{(n)} f(x_k^{(n)}), \quad n = 1, \dots,$$

where ρ is a weight function on $[a, b]$. Suppose

$$(1) \quad \lim_{n \rightarrow \infty} \sum_{k=1}^n A_k^{(n)} p(x_k^{(n)}) = \int_a^b \rho(x)p(x) dx \quad \forall p \in \mathcal{P}, \quad (3.25)$$

$$(2) \quad \sup_{n \geq 1} \sum_{k=1}^n |A_k^{(n)}| < \infty. \quad (3.26)$$

Then

$$\lim_{n \rightarrow \infty} \sum_{k=1}^n A_k^{(n)} f(x_k^{(n)}) = \int_a^b \rho(x)f(x) dx \quad \forall f \in C[a, b].$$

Proof. Denote for each integer $n \geq 1$ and $f \in C[a, b]$

$$I_n(f) = \sum_{k=1}^n A_k^{(n)} f(x_k^{(n)}).$$

Clearly, each $I_n : C[a, b] \rightarrow \mathbb{R}$ is linear. Moreover, setting

$$M = \sup_{n \geq 1} \sum_{k=1}^n |A_k^{(n)}|,$$

we have

$$|I_n(f)| \leq \sum_{k=1}^n |A_k^{(n)}| \cdot |f(x_k^{(n)})| \leq M \|f\|_{C[a,b]} \quad \forall f \in C[a,b] \quad \forall n \geq 1.$$

Now, let $f \in C[a,b]$ and let $\varepsilon > 0$. By the First Weierstrass Approximation Theorem, there exists $p \in \mathcal{P}$ such that

$$\|f - p\|_{C[a,b]} < \frac{\varepsilon}{2 \left(M + \int_a^b \rho(x) dx \right)}.$$

By the first assumption (3.25), there exists an integer $N \geq 1$ such that

$$\left| I_n(p) - \int_a^b \rho(x)p(x) dx \right| < \frac{\varepsilon}{2} \quad \forall n \geq N.$$

Therefore, for any $n \geq N$,

$$\begin{aligned} & \left| I_n(f) - \int_a^b \rho(x)f(x) dx \right| \\ & \leq |I_n(f - p)| + \left| I_n(p) - \int_a^b \rho(x)p(x) dx \right| + \int_a^b \rho(x)|p(x) - f(x)| dx \\ & < M \|f - p\|_{C[a,b]} + \frac{\varepsilon}{2} + \|p - f\|_{C[a,b]} \int_a^b \rho(x) dx \\ & < \varepsilon. \end{aligned}$$

This completes the proof. □

Colloary 3.17. *Given a sequence of interpolatory numerical quadrature*

$$\int_a^b \rho(x)f(x) dx \approx \sum_{k=1}^n A_k^{(n)} f(x_k^{(n)}), \quad n = 1, \dots$$

Suppose all the coefficients $A_k^{(n)}$ ($k = 1, \dots, n; n = 1, \dots$) are positive. Then

$$\lim_{n \rightarrow \infty} \sum_{k=1}^n A_k^{(n)} f(x_k^{(n)}) = \int_a^b \rho(x)f(x) dx \quad \forall f \in C[a,b]. \quad (3.27)$$

Proof. Let $p \in \mathcal{P}$. Then there exists an integer $N \geq 1$ such that $p \in \mathcal{P}_N$. Since an interpolatory quadrature with n quadrature points has the degree of precision $\geq n - 1$. Therefore,

$$\sum_{k=1}^n A_k^{(n)} p(x_k^{(n)}) = \int_a^b \rho(x)p(x) dx \quad \forall n \geq N.$$

This shows that the assumption (3.25) in the above theorem holds true. We have also for all $n \geq 1$ that

$$\int_a^b \rho(x) dx = \sum_{k=1}^n A_k^{(n)} = \sum_{k=1}^n |A_k^{(n)}|,$$

where we used the fact that all $A_k^{(n)}$ ($k = 1, \dots, n; n = 1, \dots$) are positive. Thus, the second assumption (3.26) in the above theorem holds true. The desired convergence (3.27) then follows from Theorem 3.16. \square

A direct consequence of this corollary and Proposition 3.14 is the following:

Colloary 3.18 (Convergence of weighted Gaussian quadrature). *For any sequence of weighted Gaussian quadrature*

$$\int_a^b \rho(x) f(x) dx \approx \sum_{k=1}^n A_k^{(n)} f(x_k^{(n)}), \quad n = 1, \dots,$$

we have

$$\lim_{n \rightarrow \infty} \sum_{k=1}^n A_k^{(n)} f(x_k^{(n)}) = \int_a^b \rho(x) f(x) dx \quad \forall f \in C[a, b]. \quad \square$$

3.7 Adaptive Method

Exercises

- Use (1) the left endpoint rectangle rule, (2) the midpoint rectangle rule, (3) the trapezoid rule, (4) Simpson's rule, and (5) the two-point Gaussian quadrature to compute the integral

$$\int_0^1 \sin x dx$$

with the number of subintervals of equal length $n = 1, \dots, 10$, respectively. Compute also all the corresponding absolute errors.

- Make a table of six columns with one for n and the other five for the the computed values by the five rules, respectively. Keep eight digits after decimal points.
 - In a single plot, display five curves showing the absolute errors in the log-log scale (i.e. $\log(\text{error})$ vs. $\log(n)$) for the five corresponding rules.
 - Discuss convergence rates for these quadrature rules based on the computational result.
- Use the trapezoid rule and Simpson's rule to compute the integral

$$\int_0^1 \sin x dx$$

with the number of subintervals of equal length $n = 2^k$, $k = 0, \dots, 8$. Then, apply the Richardson extrapolation procedure to the computed values with the trapezoid rule for the pairs $(2k - 1, 2k)$, $k = 1, \dots, 8$. Compute all the corresponding absolute errors.

- (a) Make a table of four columns with one for n , one for the computed values by the trapezoid rule, one for that by the Richardson extrapolation, and one for that by Simpson's rule. Keep eight digits after decimal points.
 - (b) In a single plot, display three curves showing in the log-log scale the absolute errors (i.e. $\log(\text{error})$ vs. $\log(n)$) for the two corresponding rules and the Richardson extrapolation.
 - (c) Discuss the computational result in terms of convergence rates.
3. Given a numerical integration formula on $[-1, 1]$

$$\int_{-1}^1 g(t) dt \approx \sum_{j=1}^n a_j g(t_j). \quad (3.28)$$

Define, for an interval $[a, b]$, $A_j = (b - a)a_j/2$ and $x_j = [(b - a)t_j + a + b]/2$, $j = 1, \dots, n$. Show that the numerical integration formula on $[a, b]$

$$\int_a^b f(x) dx \approx \sum_{j=1}^n A_j f(x_j)$$

has the same degree of precision as that of the formula (3.28).

4. Find A, B, C such that the weighted numerical quadrature

$$\int_{-2}^2 |x|f(x) dx \approx Af(-1) + Bf(0) + Cf(1)$$

is exact for polynomials of degree as high as possible. What is the degree of precision of the quadrature?

5. Let $h > 0$. Find A, B, C, D so that the numerical quadrature

$$\int_{-h}^h f(x) dx \approx Af(-h) + Bf(0) + Cf(h) + Dhf'(h)$$

is exact for polynomials of degree as high as possible. What is the degree of precision of the quadrature?

6. Find A, B, C, D such that the numerical quadrature

$$\int_0^1 f(x) dx \approx Af(0) + Bf(1) + Cf''(0) + Df''(1)$$

is exact for polynomials of degree as high as possible. What is the degree of precision of the quadrature?

7. Let $n \geq 1$ be an integer and $a \leq x_1 < \cdots < x_n \leq b$. Suppose that the degree of precision of the numerical integration formula

$$\int_a^b f(x) dx \approx \sum_{k=1}^n A_k f(x_k)$$

is $\geq n - 1$. Define for each integer $j \geq 0$

$$F_j(t) = \int_a^b (x-t)_+^j dx - \sum_{k=1}^n A_k (x_k - t)_+^j.$$

Show that

$$\int_a^b F_j(t) dt = 0 \quad j = 0, \dots, n-1.$$

8. Consider the trapezoidal formula

$$\int_a^b f(x) dx \approx \frac{1}{2}(b-a)[f(a) + f(b)].$$

- (a) Show that the degree of precision of the formula is $m = 1$.
 (b) Calculate explicitly the Peano kernel K_1 of the formula and show that the kernel does change sign in $[a, b]$.
 (c) Let $f \in C^2[a, b]$. Show that there exists $\xi \in (a, b)$ such that

$$\int_a^b f(x) dx - \frac{1}{2}(b-a)[f(a) + f(b)] = -\frac{1}{12}(b-a)^3 f''(\xi).$$

- (d) Let $N \geq 1$ be an integer, $h = (b-a)/N$, and $x_j = a + jh$, $j = 0, \dots, N$. Prove for $f \in C^2[a, b]$ the error formula for the composite trapezoidal formula

$$\int_a^b f(x) dx - \left\{ \frac{h}{2}[f(a) + f(b)] + h \sum_{j=1}^{N-1} f(x_j) \right\} = -\frac{(b-a)f''(\eta)}{12}h^2,$$

where $\eta \in (a, b)$ depends on f .

9. Find an integer $N \geq 1$, as small as possible, so that

$$\left| \int_0^1 e^x dx - T_N \right| \leq 10^{-12},$$

where T_N is the numerical integration value (without round-off error) of the function e^x over $[0, 1]$ using the composite trapezoidal rule with N subintervals of equal length.

10. Let $p_3 \in \mathcal{P}_3$ be the Hermite interpolation polynomial of $f \in C^1[a, b]$ determined by

$$p_3(a) = f(a), \quad p_3'(a) = f'(a), \quad p_3(b) = f(b), \quad p_3'(b) = f'(b).$$

(a) Show that

$$\int_a^b p_3(x) dx = \frac{1}{2}(b-a)[f(a) + f(b)] - \frac{1}{12}(b-a)^2[f'(b) - f'(a)].$$

(b) Determine the degree of precision of the numerical quadrature

$$\int_a^b f(x) dx \approx \frac{1}{2}(b-a)[f(a) + f(b)] - \frac{1}{12}(b-a)^2[f'(b) - f'(a)]. \quad (3.29)$$

(c) Calculate explicitly the Peano kernel of the numerical quadrature (3.29).

(d) Derive the error formula for the numerical quadrature (3.29).

(e) Let $N \geq 1$ be an integer, $h = (b-a)/N$, and $x_j = a + jh$, $j = 0, \dots, N$. Derive the composite integration formula based on the formula (3.29). Show that the composite formula is the same as the composite trapezoidal formula for functions $f \in C^1[a, b]$ such that $f'(a) = f'(b)$.

11. Let $p \in \mathcal{P}_5$ be the Hermite interpolation polynomial of $f \in C^1[-1, 1]$ determined by

$$p(x_j) = f(x_j) \quad \text{and} \quad p'(x_j) = f'(x_j), \quad j = 0, 1, 2,$$

where $x_0 = -1, x_1 = 0, x_2 = 1$.

(a) Show that

$$\int_{-1}^1 p(x) dx = \frac{1}{15} [7f(-1) + 16f(0) + 7f(1) + f'(-1) - f'(1)].$$

(b) Show that the degree of precision of the numerical integration formula

$$\int_{-1}^1 f(x) dx \approx \frac{1}{15} [7f(-1) + 16f(0) + 7f(1) + f'(-1) - f'(1)]. \quad (3.30)$$

is $m = 5$.

(c) Derive the error formula for the numerical integration formula (3.30).

(d) Derive the composite numerical integration formula corresponding to the formula (3.30).

12. Let $a \leq x_0 < \dots < x_n \leq b$. Show that there exist $n + 1$ real numbers $\gamma_0, \dots, \gamma_n$ such that

$$\int_a^b p(x) dx = \sum_{j=0}^n \gamma_j p(x_j) \quad \forall p \in \mathcal{P}_n.$$

13. Consider the Newton-Cotes formula

$$\int_a^b f(x) dx \approx \sum_{j=0}^n A_j f(x_j)$$

with $n + 1$ points $x_j = a + j(b-a)/n$, $j = 0, \dots, n$.

- (a) Show that $A_j = A_{n-j}$ for $j = 0, \dots, [n/2]$.
 (b) Show by direct calculation that the degree of precision of the formula is n if n is odd and is $n + 1$ if n is even.
14. Let $n \geq 2$ be an even number, $\omega_n(x) = \prod_{j=0}^n (x - j)$, and

$$\Omega_n(x) = \int_0^x \omega_n(t) dt.$$

- Show that $\Omega_n(0) = \Omega_n(n) = 0$ and that $\Omega_n(x) > 0$ for all $x \in (0, n)$.
15. Let x_1, \dots, x_n be n distinct points in $[a, b]$ and A_1, \dots, A_n be n real numbers. If the degree of precision of the weighted numerical integration formula

$$\int_a^b \rho(x) f(x) dx \approx \sum_{j=1}^n A_j f(x_j)$$

- with ρ a weight function on $[a, b]$ is $2n - 1$, then it must be the weighted Gaussian formula on $[a, b]$ with the weight function ρ .
16. Let $n \geq 2$ be an integer and x_1, \dots, x_n the n distinct roots in $(-1, 1)$ of the n th Legendre polynomial P_n . Set

$$l_j(x) = \frac{P_n(x)}{(x - x_j)P_n'(x_j)} \quad \text{and} \quad A_j = \int_{-1}^1 l_j(x) dx, \quad j = 1, \dots, n.$$

- (a) Show that
- $$\int_{-1}^1 p(x)q(x) dx = \sum_{j=1}^n A_j p(x_j)q(x_j) \quad \forall p, q \in \mathcal{P}_{n-1}.$$
- (b) Show that
- $$A_j = \int_{-1}^1 [l_j(x)]^2 dx > 0, \quad j = 1, \dots, n.$$
17. Let $\{Q_n\}_{n=0}^\infty$ be a system of orthogonal polynomials with each $\deg Q_n = n$ with respect to the inner product in $L_\rho^2[a, b]$, where ρ is a weight function on $[a, b]$. Fix $n \geq 1$. Let x_1, \dots, x_n be the n distinct roots of Q_n in (a, b) . Let

$$\int_a^b \rho(x) f(x) dx \approx \sum_{j=1}^n A_j f(x_j)$$

be the corresponding weighted Gaussian quadrature. Show that

$$\sum_{j=1}^n A_k Q_k(x_j) = 0, \quad k = 1, \dots, 2n - 1.$$

18. (Gautschi) Consider a weighted Gaussian formula

$$\int_a^b \rho(x)f(x) dx \approx \sum_{j=1}^n A_j f(x_j)$$

with ρ a weight function on $[a, b]$. Show that for any $f \in C[a, b]$ the error

$$e_n(f) = \int_a^b \rho(x)f(x) dx - \sum_{j=1}^n A_j f(x_j)$$

satisfies

$$|e_n(f)| \leq 2 \left(\int_a^b \rho(x) dx \right) E_{2n-1}(f),$$

where

$$E_{2n-1}(f) = \min_{q \in \mathcal{P}_{2n-1}} \|f - q\|_{C[a,b]}.$$

19. Let $Q_n \in \mathcal{P}_n$ be the n th orthogonal polynomial with respect to the weight ρ on $[a, b]$, $n = 0, \dots$. Fix $n \geq 1$. Let x_1, \dots, x_n be the n distinct roots of Q_n in (a, b) . Let

$$\int_a^b \rho(x)f(x) dx \approx \sum_{j=1}^n A_j f(x_j)$$

be the corresponding weighted Gaussian quadrature. Show that

$$A_j = \frac{a_n}{a_{n-1} Q'_n(x_j) Q_{n-1}(x_j)}, \quad j = 1, \dots, n,$$

where a_k is the leading coefficient of Q_k ($k = 0, \dots$).

20. Let $n \geq 1$ be an integer. The Gauss-Chebyshev quadrature is the weighted Gaussian quadrature on $[-1, 1]$ with the weight $1/\sqrt{1-x^2}$ using $x_j = \cos(2j-1)\pi/2n$ ($j = 1, \dots, n$), the n roots of the n th Chebyshev polynomial $T_n(x) = \cos(n \arccos x)$. Show that the Gauss-Chebyshev formula is given by

$$\int_{-1}^1 \frac{f(x)}{\sqrt{1-x^2}} dx \approx \frac{\pi}{n} \sum_{j=1}^n f(x_j).$$

21. (a) Show for any $f \in C[0, 1]$ that

$$\int_0^1 B_n f(x) dx = \frac{1}{n+1} \sum_{k=0}^n f\left(\frac{k}{n}\right), \quad n = 0, \dots,$$

where $B_n f$ ($n = 0, \dots$) are the Bernstein polynomials of f .

(b) Show that the degree of precision of the numerical integration formula

$$\int_0^1 f(x) dx \approx \frac{1}{n+1} \sum_{k=0}^n f\left(\frac{k}{n}\right)$$

is $m = 1$ for all $n = 0, \dots$

(c) Show that

$$\lim_{n \rightarrow \infty} \frac{1}{n+1} \sum_{k=0}^n f\left(\frac{k}{n}\right) = \int_0^1 f(x) dx \quad \forall f \in C[0, 1].$$

22. (Bernoulli polynomials) Define

$$B_0(x) = 1, \quad B_1(x) = x - \frac{1}{2},$$

$$B'_{n+1}(x) = (n+1)B_n(x) \quad \text{and} \quad \int_0^1 B_{n+1}(x) dx = 0, \quad n = 2, \dots$$

(a) Prove that, for each $n \geq 0$, B_n is a polynomial of degree n with leading coefficient 1.

(b) Prove the identities

$$\begin{aligned} B_n(x+1) - B_n(x) &= nx^{n-1}, & n = 0, \dots, \\ B_n(1-x) &= (-1)^n B_n(x), & n = 0, \dots \end{aligned}$$

(c) Prove the identities

$$\begin{aligned} B_n(0) &= B_n(1), & n = 2, \dots, \\ B_{2n+1}(0) &= 0, & n = 1, \dots \end{aligned}$$

23. Prove

$$\sum_{k=1}^n k^3 = \left(\frac{n(n+1)}{2}\right)^2$$

by the Euler-Maclaurin summation formula.

24. Let $f \in C[a, b]$ and denote by $I(f)$ the integral of f over $[a, b]$. Let $N \geq 1$ be an integer, $h = (b-a)/2N$, and $x_j = a + jh$, $j = 0, \dots, 2N$. Let T_N , T_{2N} , and S_N denote, respectively, the approximate value of $I(f)$ by the composite trapezoidal rule with N subintervals $[x_{2j-1}, x_{2j}]$, $j = 1, \dots, N$, by the composite trapezoidal rule with $2N$ subintervals $[x_{j-1}, x_j]$, $j = 0, \dots, 2N$, and by the composite Simpson rule with N subintervals $[x_{2j-1}, x_{2j}]$, $j = 1, \dots, N$. Prove that the Richardson extrapolation using T_N and T_{2N} leads to exactly S_N , i.e.,

$$S_N = \frac{4T_{2N} - T_N}{3}.$$

Bibliography

- [1] E. W. Cheney. *Introduction to Approximation Theory*. McGraw-Hill, 1966.
- [2] P. J. Davis. *Interpolation and Approximation*. Blaisdell, 1963.
- [3] P. J. Davis and P. H. Rabinowitz. *Methods of Numerical Integration*. Academic press, 1975.
- [4] H. Engles. *Numerical Quadrature and Cubature*. Academic Press, 1980.
- [5] E. I. Isaacson and H. B. Keller. *Analysis of Numerical Methods*. Dover, 1994.
- [6] P. P. Korovkin. *Linear Operators and Approximation Theory*. Hindustan, 1960.
- [7] P. M. J. D. Powell. *Approximation Theory and Methods*. Cambridge University Press, 1981.
- [8] T. J. Rivlin. *Introduction to the Approximation of Functions*. Blaisdell, 1969 (Dover, 1987).
- [9] G. Szegő. *Orthogonal Polynomials*. Amer. Math. Soc., 1959.
- [10] J. Todd. *Introduction to the Constructive Theory of Functions*. Academic Press, 1963.
- [11] L.-Z. Xu, R.-H. Wang, and Y.-S. Zhou. *Theory and Methods of Function Approximation*. Shanghai Sci. Tech. Press, 1983 (in Chinese).