Contents

1 Polynomial Approximation **5**
- 1.1 The Weierstrass Theorem . . . 5
- 1.2 Best Uniform Approximation . . . 8
- 1.3 Chebyshev Polynomials . . . 14
- 1.4 Least-Squares Approximation . . . 17
- 1.5 Orthogonal Polynomials . . . 21
- 1.6 More Properties of Orthogonal Polynomials . . . 24
- 1.7 Legendre Polynomials . . . 29
- Exercises . . . 33

2 Polynomial Interpolation **41**
- 2.1 Lagrange Interpolation . . . 41
- 2.2 Newton's Formula and Divided Differences . . . 46
- 2.3 Peano Kernal and the Remainder Theorem . . . 52
- 2.4 Hermite Interpolation and Divided Differences with Repeated Points . . . 56
- 2.5 Convergence of Interpolation Polynomials . . . 58
- 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 Weighted Gaussian Quadrature . . . 82
- 3.5 Convergence of Sequences of Numerical Quadrature . . . 88
- 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*

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

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

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

where

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

For each $n \geq 0$, we call $B_n f$ the *nth 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 $n \geq 0$, $B_n : C[0, 1] \to \mathcal{P}_n$ is linear, i.e.,*

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

(2) *For each* $n \geq 0$, $B_n : C[0,1] \to \mathcal{P}_n$ *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,*

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

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

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

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

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

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

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

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,

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

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

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

We claim that

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

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

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

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

Since

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

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

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

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

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

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

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

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

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

This is (1.1). □

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

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

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

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

$$|\,\|f\| - \|g\|\,| \le \|f-g\| \qquad \forall f, g \in C[a,b]. \tag{1.11}$$

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 $p^{(k)} \to f$ in $C[a,b]$, i.e., $\|f - p^{(k)}\| \to 0$ as $k \to \infty$.

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 $n \ge 1$, let $L_n : C[a,b] \to \mathcal{P}_n$ be an operator. Assume:*

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

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

1.2 Best Uniform Approximation

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

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

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

Proposition 1.4. *We have*

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

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*

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

Since (1.13) can be written as

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

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 $n \ge 0$, there exists a best uniform approximation of f in $\mathcal{P}_n$.*

Proof. Let $f \in C[a,b]$ and fix an integer $n \ge 0$. 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 $F : \mathbb{R}^{n+1} \to \mathbb{R}$ by

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

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

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

Step 1. We show that the function $F : \mathbb{R}^{n+1} \to \mathbb{R}$ satisfies the following properties:
(1) $F : \mathbb{R}^{n+1} \to \mathbb{R}$ is continuous;
(2) $F(c) \ge 0$ for any $c \in \mathbb{R}^{n+1}$;
(3) $F(c) \to +\infty$ as $c \to \infty$.
By (1.11), we have

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

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

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

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

where

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

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 $G : \mathbb{R}^{n+1} \to \mathbb{R}$ then implies that

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

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

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

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

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

proving Part (3).

Step 2. Let $m = \inf_{c \in \mathbb{R}^{n+1}} F(c) \geq 0$. Since $\lim_{c \to \infty} F(c) = +\infty$, 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 $\mathbb{R}^{n+1}$. □

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 $a \leq x_1 < \cdots < x_{n+2} \leq b$ such that*

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

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.,

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

Equivalently,

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

Definition 1.8 (Change of sign)**.** *A function $g : (a, b) \to \mathbb{R}$ 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) *$g(x) \ge 0$ for all $x \in (z - \varepsilon, z)$ and $g(x) \le 0$ for all $x \in (z, z + \varepsilon)$;*
(2) *both one-sided limits $g(z-)$ and $g(z+)$ exist and $g(z-) \ne g(z+)$,*

or the following hold true:

(1) *$g(x) \le 0$ for all $x \in (z - \varepsilon, z)$ and $g(x) \ge 0$ for all $x \in (z, z + \varepsilon)$;*
(2) *both one-sided limits $g(z-)$ and $g(z+)$ exist and $g(z-) \ne g(z+)$.*

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 p_n of f, then there would exit $p \in \mathcal{P}_n$ such that

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

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

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

It follows from (1.16) that

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

Consequently,

$$\operatorname{sign} p(x_k) = \operatorname{sign} (p(x_k) - f(x_k) + f(x_k)) = \operatorname{sign} f(x_k), \qquad k = 1, \ldots, n + 2,$$

where $\operatorname{sign} z = z/|z|$ for any nonzero $z \in \mathbb{R}$ and $\operatorname{sign} 0 = 0$. 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 $0 \le m \le n$ 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, $1 \le m \le n$. 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 $\operatorname{sign} f$ alternates at $m+1$ points in Z, there exist ξ_k $(1 \le k \le m)$ with $a \le \xi_1 < \cdots < \xi_m \le b$ such that $\operatorname{sign} f$ alternates in $Z \cap [\xi_k, \xi_{k+1}]$ $(k = 1, \ldots, m)$. Define

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

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

$$\operatorname{sign} p(x) = \operatorname{sign} f(x) \qquad \forall x \in Z. \tag{1.17}$$

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

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

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

Step 2. We show that

$$\left\| f - \frac{\varepsilon}{\|p\|} p \right\| < \|f\| \tag{1.19}$$

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

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

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

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

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

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

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

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

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

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

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

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

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 $(k = 1, \ldots, n + 2)$ with $a \leq x_1 < \cdots < x_{n+2} \leq b$ such that

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

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

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

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

1.3 Chebyshev Polynomials

Consider $n \ge 1$ 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

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

Denote for any integer $k \ge 0$

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

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

$$\|\widetilde{T}_n\|_{C[-1,1]} = \min_{\widetilde{p} \in \widetilde{\mathcal{P}}_n} \|\widetilde{p}\|_{C[-1,1]}. \tag{1.22}$$

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

To find $\widetilde{T}_n$, we consider for each $n \ge 0$[1]

$$T_n(x) = \cos(n \arccos x) \qquad \forall x \in [-1, 1]. \tag{1.23}$$

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

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

Since

$$\cos(n+1)\theta + \cos(n-1)\theta = 2\cos\theta \cos n\theta, \tag{1.24}$$

we obtain

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

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 $n \ge 1$ the leading coefficient of T_n is 2^{n-1}.

[1] See [**?**] for more on the derivation of $\widetilde{T}_n$.

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

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

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 $n \geq 1$,

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

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.

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

(2) Recurrence.

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

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.

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

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

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

(5) Differential equation.

$$(1-x^2)T_n''(x) - xT_n'(x) + n^2 T_n(x) = 0, \qquad n = 0, \ldots.$$

(6) The generating function.

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

(7) Rodrigues' formula.

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

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

(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

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

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 1.22 in Section 1.6 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

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

Since

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

we have

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

completing the proof. □

1.4 Least-Squares Approximation

Definition 1.11 (Weight functions)**.** *A* weight function *on a finite intervale* $[a,b]$ *is a non-negative, integrable function* $\rho : (a,b) \to \mathbb{R}$ *that satisfies*

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

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 $\rho : [a,b] \to \mathbb{R}$ 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 $L^2_\rho(a,b)$ the set of all measurable functions $f : [a,b] \to \mathbb{R}$ 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 $L^2_\rho(a,b)$. Under the usual addition and scalar multiplication, $L^2_\rho(a,b)$ is a vector space. Clearly, $\mathcal{P} \subset C[a,b] \subseteq$

$L^2_\rho(a,b)$. If $f, g \in L^2_\rho(a,b)$, then

$$\int_a^b \rho(x)|f(x)g(x)|\,dx \le \int_a^b \rho(x)\frac{[f(x)]^2+[g(x)]^2}{2}dx < \infty. \tag{1.28}$$

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

Proposition 1.12 (The Cauchy-Schwarz inequality)**.** *We have*

$$\left|\int_a^b \rho(x)f(x)g(x)\,dx\right| \le \sqrt{\int_a^b \rho(x)[f(x)]^2dx}\sqrt{\int_a^b \rho(x)[g(x)]^2dx} \qquad \forall f, g \in L^2_\rho(a,b). \tag{1.29}$$

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

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

If the leading coefficient of ϕ is 0, then the inequality (1.29) holds true trivially. If it is nonzero, then the discriminant of ϕ is non-positive, i.e.,

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

This leads to the inequality in (1.29). □

Definition 1.13 (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*

$$\int_a^b \rho(x)[f(x)-p_n(x)]^2dx \le \int_a^b \rho(x)[f(x)-q_n(x)]^2dx \qquad \forall q_n \in \mathcal{P}_n. \tag{1.30}$$

Theorem 1.14 (Existence, uniqueness, and characterization of least-squares approximations)**.** *Let $f \in L^2_\rho(a,b)$. Let $n \ge 0$ 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*

$$\int_a^b \rho(x)[f(x)-p_n(x)]q_n(x)dx = 0 \qquad \forall q_n \in \mathcal{P}_n. \tag{1.31}$$

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

$$\int_a^b \rho(x)[f(x)-p_n(x)]^2dx = \int_a^b \rho(x)[f(x)]^2dx - \int_a^b \rho(x)[p_n(x)]^2dx. \tag{1.32}$$

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

$$G_{n+1} := \left[\int_a^b \rho(x)x^{i+j}dx\right]_{i,j=0}^{n}. \tag{1.33}$$

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

Lemma 1.15. *The matrix G_{n+1} is symmetric positive definite.*

Proof. Define

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

Clearly, $Q:\mathbb{R}^{n+1}\to\mathbb{R}$ is a non-negative quadratic form. If $Q(\xi)=0$ for some $\xi\in\mathbb{R}^{n+1}$, then, by Lemma 1.16, $\sum_{i=0}^{n}\xi_i x^i=0$ for all $x\in[a,b]$. Hence, $\xi=0$ in $\mathbb{R}^{n+1}$. Therefore, $Q:\mathbb{R}^{n+1}\to\mathbb{R}$ is a positive quadratic form.

It is clear that

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

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

Proof of Theorem 1.14. (1) Define

$$F(c) = \int_a^b \rho(x)\left[f(x)-\sum_{i=0}^{n}c_i x^i\right]^2 dx \qquad \forall c=(c_0,\ldots,c_n)\in\mathbb{R}^{n+1}. \tag{1.34}$$

Clearly, $F:\mathbb{R}^{n+1}\to\mathbb{R}$ 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 (1.33). By Lemma 1.15, the Hessian matrix of F is therefore symmetric positive definite. Consequently, $F:\mathbb{R}^{n+1}\to\mathbb{R}$ is a convex quadratic form. Since it is non-negative, it admits a unique minimizer $\hat{c}=(\hat{c}_0,\ldots,\hat{c}_n)\in\mathbb{R}^{n+1}$. Define

$$p_n(x) = \sum_{i=0}^{n}\hat{c}_i x^i \qquad \forall x\in\mathbb{R}. \tag{1.35}$$

Then, $p_n \in \mathcal{P}_n$ is the unique polynomial in $\mathcal{P}_n$ that satisfies (1.30).

(2) From Part (1), the unique least-squares approximation $p_n \in \mathcal{P}_n$ is given by (1.35), where $\hat{c} = (\hat{c}_0, \ldots, \hat{c}_n) \in \mathbb{R}^{n+1}$ is the unique minimizer of the convex quadratic function $F : \mathbb{R}^{n+1} \to \mathbb{R}$. Clearly, $\hat{c}$ is the unique critical point of F determined by

$$\partial_j F(\hat{c}) = 0, \qquad j = 0, \ldots, n. \tag{1.36}$$

By the definition of $F : \mathbb{R}^{n+1} \to \mathbb{R}$ (cf. (1.34)) and the Chain Rule, we obtain for any $c = (c_0, \ldots, c_n) \in \mathbb{R}^{n+1}$ that

$$\partial_j F(c) = 2 \int_a^b \rho(x) \left[f(x) - \sum_{i=0}^{n} c_i x^i \right] \left(-x^j \right) dx, \qquad j = 0, \ldots, n. \tag{1.37}$$

Therefore, by the definition of p_n in (1.35), the system of equations (1.36) is equivalent to

$$\int_a^b \rho(x) \left[f(x) - p_n(x) \right] x^j dx, \qquad j = 0, \ldots, n,$$

which in turn is equivalent to (1.31), since $\mathcal{P}_n = \text{Span}\{1, x, \ldots, x^n\}$.

(3) Suppose $p_n \in \mathcal{P}_n$ is the least-squares approximation of f. By (1.31) with $q_n = p_n$, we have

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

Therefore,

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

This implies (1.32). □

From the proof, cf. (1.36) and (1.37), 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

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

The coefficient matrix of this system of linear equations is exactly G_{n+1}, defined in (1.33).

1.5 Orthogonal Polynomials

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

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

By (1.28), $\langle f, g\rangle$ is finite for any $f, g \in L^2_\rho(a, b)$. The mapping $\langle \cdot, \cdot\rangle : L^2_\rho(a, b) \times L^2_\rho(a, b) \to \mathbb{R}$ clearly satisfies the following properties:

(1) Symmetry.

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

(2) Bilinearity.

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

(3) Non-negativity.

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

Lemma 1.16. *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 $\langle \cdot, \cdot\rangle : C[a, b] \times C[a, b] \to \mathbb{R}$ is in fact an inner product of $C[a, b]$, i.e., it is symmetric, bilinear, and positive. Here, the positivity means

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

The "only if" part in the last equation follows from Lemma 1.16. The norm of $C[a,b]$ induced by this inner product is

$$\|f\|_{L^2_\rho(a,b)} = \sqrt{\langle f, f\rangle} = \sqrt{\int_a^b \rho(x)[f(x)]^2 dx} \tag{1.38}$$

for any $f \in C[a,b]$. This is called the $L^2_\rho(a,b)$-norm. We will also use (1.38) 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 1.17 (Orthogonal polynomials)**.** *A sequence of polynomials Q_n $(n = 0, ...)$ 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 $(n = 0, ...)$ 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

$$\langle Q_i, Q_j\rangle = \delta_{i,j} \qquad \forall i, j \geq 0, \tag{1.39}$$

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 1.7.

Lemma 1.18. *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) $\mathcal{P}_n = Span\{Q_0, \dots, Q_n\}$. *Moreover,*

$$p_n = \sum_{k=0}^{n} \frac{\langle p_n, Q_k\rangle}{\langle Q_k, Q_k\rangle} Q_k \qquad \forall p_n \in \mathcal{P}_n. \tag{1.40}$$

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 (1.39) 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 1.16. 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 $p_n \in \mathcal{P}_n = \text{Span}\{Q_0, \ldots, Q_n\}$. There exist constants $c_i \in \mathbb{R}$ $(i = 0, \ldots, n)$ such that $p_n = \sum_{k=0}^{n} c_i Q_i$. Multiplying both sides of this equation by ρQ_j for an arbitrary but fixed j with $0 \le j \le n$ and then integrate over $[a, b]$, we obtain by (1.39) that $\langle p_n, Q_j \rangle = c_j \langle Q_j, Q_j \rangle$. This implies (1.40). □

Theorem 1.19 (Characterization of least-squares approximations using orthonormal polynomials)**.** *Let $Q_0, \ldots, Q_n$ 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*

$$p_n = \sum_{k=0}^{n} \langle f, Q_k \rangle Q_k. \tag{1.41}$$

Moreover,

$$\|f - q_n\|^2 = \|f\|^2 - \sum_{k=0}^{n} \langle f, Q_k \rangle^2. \tag{1.42}$$

Proof. Clearly, the polynomial p_n defined in (1.41) is in $\mathcal{P}_n$. Fix an arbitrary integer j with $0 \le j \le n$. Since $Q_0, \ldots, Q_n$ 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 1.18,

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

This is exactly (1.31). Therefore, by Theorem 1.14, $p_n \in \mathcal{P}_n$ is the unique least-squares approximation of f in $\mathcal{P}_n$.

It follows from (1.41), the symmetry and bilinearity of $\langle \cdot, \cdot \rangle$, and the fact that $Q_0, \ldots, Q_n$ are orthonormal that

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

This, together with Part (3) of Theorem 1.14, implies (1.42). □

Theorem 1.20 (The Gram-Schmidt orthogonalization)**.** *Consider $L^2_\rho(a, b)$. Let $f_0(x) = 1$ and $f_k(x) = x^k$ for any integer $k \ge 1$. Define*

$$g_0(x) = \frac{1}{\sqrt{\langle f_0, f_0 \rangle}} f_0(x), \tag{1.43}$$

$$\begin{cases} \hat{g}_k(x) = f_k(x) - \displaystyle\sum_{j=1}^{k-1} \langle f_k, g_j \rangle g_j(x), & \\ g_k(x) = \dfrac{1}{\sqrt{\langle \hat{g}_k, \hat{g}_k \rangle}} \hat{g}_k(x), & k = 2, \ldots \end{cases} \tag{1.44}$$

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

Proof. Clearly, g_0 and g_1 are polynomial of degrees 0 and 1, respectively. Moreover, direct calculations using (1.43) and (1.44) 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

$$\langle g_i, g_j\rangle = \delta_{ij}, \qquad 0 \leq i, j \leq k. \tag{1.45}$$

Since f_{k+1} is a polynomial of degree $k+1$, we see from (1.44) 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 (1.44) and (1.45) 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

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

Therefore, (1.45) is true with k replaced by $k+1$. Consequently, g_k $(k = 0, \ldots)$ are orthonormal polynomials. □

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

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

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

1.6 More Properties of Orthogonal Polynomials

Theorem 1.21 (The three-term recurrence)**.** *Consider* $L^2_\rho(a, b)$.

(1) *Define*

$$\begin{aligned}
\widetilde{Q}_0(x) &= 1,\\
\widetilde{Q}_1(x) &= x - a_1,\\
\widetilde{Q}_n(x) &= (x - a_n)\widetilde{Q}_{n-1}(x) - b_n\widetilde{Q}_{n-2}(x), \qquad n = 2, \dots,\\
a_n &= \frac{\langle x\widetilde{Q}_{n-1}, \widetilde{Q}_{n-1}\rangle}{\langle \widetilde{Q}_{n-1}, \widetilde{Q}_{n-1}\rangle}, \qquad n = 1, \dots,\\
b_n &= \frac{\langle \widetilde{Q}_{n-1}, \widetilde{Q}_{n-1}\rangle}{\langle \widetilde{Q}_{n-2}, \widetilde{Q}_{n-2}\rangle}, \qquad n = 2, \dots.
\end{aligned}$$

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

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

$$Q_n(x) = (A_n x + B_n)Q_{n-1}(x) - C_n Q_{n-2}(x) \qquad n = 2, \dots. \tag{1.46}$$

where

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

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

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

we have

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

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

$$\langle \widetilde{Q}_j, \widetilde{Q}_k\rangle = 0, \qquad \text{if } 0 \leq j, k \leq n-1 \text{ and } j \neq k. \tag{1.47}$$

Clearly, $\widetilde{Q}_n \in \widetilde{\mathcal{P}}_n$. We have by the definition of $\widetilde{Q}_n$ and (1.47) that

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

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

where the last equation follows from that fact that $x\widetilde{Q}_{n-2} - \widetilde{Q}_{n-1} \in \mathcal{P}_{n-2}$. Let $0 \le k \le n-3$ with $n \ge 3$. We have again by the definition of $\widetilde{Q}_n$ and (1.47) that

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

since $x\widetilde{Q}_k \in \mathcal{P}_{k+1} \subseteq \mathcal{P}_{n-2}$. Therefore, (1.47) holds true with $n-1$ replaced by n. Part (1) is thus proved.

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

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

This is (1.46). □

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

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

$$b_2 = \frac{\langle \widetilde{Q}_1, \widetilde{Q}_1 \rangle}{\langle \widetilde{Q}_0, \widetilde{Q}_0 \rangle} = \frac{\int_{-1}^{1} x^2 dx}{\int_{-1}^{1} 1 dx} = \frac{1}{3},$$
$$\widetilde{Q}_2(x) = (x - a_2)\widetilde{Q}_1(x) - b_2 \widetilde{Q}_0(x) = x^2 - \frac{1}{3}.$$

Thus, $\widetilde{Q}_0, \widetilde{Q}_1, \widetilde{Q}_2$ 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 1.22 (Minimization)**.** *Let Q_n $(n = 0, \dots)$ be orthogonal polynomials in $L^2_\rho(a,b)$. Suppose $n \geq 1$ and $Q_n \in \widetilde{\mathcal{P}}_n$, then Q_n is the unique polynomial in $\widetilde{\mathcal{P}}_n$ that satisfies*

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

Proof. Since $Q_n \in \widetilde{\mathcal{P}}_n$, 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,

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

Theorem 1.14 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. □

Colloary 1.23 (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 1.22, $(1/\alpha_n)P_n = (1/\beta_n)Q_n$ 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 $n \geq 0$. □

Theorem 1.24 (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 $a < x_1 < \cdots < x_k < b$. 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). □

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

$$K_n(x, t) = \sum_{k=0}^{n} Q_k(x) Q_k(t). \tag{1.48}$$

We call $K_n : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ 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 (1.41) in Theorem 1.19 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 1.25 (The Christoffel-Darboux identity)**.** *Let Q_n $(n = 0, \dots)$ be orthonormal polynomials in $L^2_\rho[a, b]$ and $K_n : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ the Dirichlet kernel associated with $Q_0, \dots, Q_n$ for each $n \geq 0$. Then,*

$$K_n(x, t) = \frac{\alpha_n}{\alpha_{n+1}} \frac{Q_{n+1}(x) Q_n(t) - Q_{n+1}(t) Q_n(x)}{x - t} \qquad \forall x, t \in \mathbb{R} \text{ with } x \neq t, \tag{1.49}$$

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

The formula (1.49) is called the *Christoffel-Darboux identity.*

Proof of Theorem 1.25. By Theorem 1.21, the orthonormal polynomials Q_n $(n = 0, \dots)$ satisfy (1.46) 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 $n \geq 1$,

$$A_{k+1}^{-1} \left[Q_{k+1}(x) Q_k(t) - Q_{k+1}(t) Q_k(x) \right]$$

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

Summing over these, we obtain that

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

since $Q_0(x) = Q_0(t) = \alpha_0$. This leads to (1.49) for $n \geq 2$. For the case $n = 0$ or 1, one can directly verify that (1.49) is true. □

1.7 Legendre Polynomials

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

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

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

$$
P_n(x) = \frac{1}{2^n n!}\frac{d^n}{dx^n}\left[\left(x^2 - 1\right)^n\right], \qquad n = 0, \ldots.
$$

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 1.26 (Properties of Legendre polynomials)**.**

(1) Orthogonality.

$$\int_{-1}^{1} P_m(x)P_n(x)\,dx = \begin{cases} 0 & \text{if } m \neq n, \\ 2/(2n+1) & \text{if } m = n. \end{cases} \tag{1.50}$$

(2) Recurrence.

$$\begin{aligned} &.P_0(x) = 1, \\ &P_1(x) = x, \\ &(n+1)P_{n+1}(x) - (2n+1)xP_n(x) + nP_{n-1}(x) = 0, \qquad n = 1, \dots. \end{aligned} \tag{1.51}$$

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

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

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

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

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

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

$$(1-x^2)P_n''(x) - 2xP_n'(x) + n(n+1)P_n(x) = 0. \tag{1.52}$$

(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*

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

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

(1) Let m, n be non-negative integers. If one of them is 0, then it is clear that (1.50) 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

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

$$
\begin{aligned}
&= -\phi_m^{(m+1)}(x)\phi_n^{(n-2)}(x)\big|_{x=-1}^{x=1} + (-1)^2 \int_{-1}^{1} \phi_m^{(m+2)}(x)\phi_n^{(n-2)}(x)\,dx \\
&= \cdots \\
&= (-1)^m \int_{-1}^{1} \phi_m^{(2m)}(x)\phi_n^{(n-m)}(x)\,dx \\
&= (-1)^m (2m)! \int_{-1}^{1} \phi_n^{(n-m)}(x)\,dx.
\end{aligned}
$$

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

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

This leads to

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

(2) Let $\widetilde{P}_n \in \widetilde{\mathcal{P}}_n$ $(n = 0, \dots)$ be the unique orthogonal polynomials each with leading coefficient 1 in $L^2(-1,1)$. By Theorem 1.21, we have

$$
\widetilde{P}_n(x) = (x - a_n)\widetilde{P}_{n-1}(x) - b_n \widetilde{P}_{n-2}(x), \qquad n = 2, \dots, \tag{1.53}
$$

where

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

By Corollary 1.23, $\widetilde{P}_n$ and P_n differ only by a constant. Comparing their leading coefficients, we have

$$\widetilde{P}_n = \frac{2^n (n!)^2}{(2n)!} P_n, \qquad n = 0, \ldots. \tag{1.54}$$

In particular, $\widetilde{P}_n \in \widetilde{\mathcal{P}}_n$ is even (odd) as P_n if n is even (odd). Hence, the function $x[\widetilde{P}_{n-1}]^2$ is an odd function, and its integral over $(-1, 1)$ vanishes. Therefore, all $a_n = 0$. By (1.54) and (1.50),

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

Consequently, we have by (1.53) with n replaced by $n+1$ that

$$\widetilde{P}_{n+1}(x) = x\widetilde{P}_n(x) - \frac{n^2}{4n^2-1} \widetilde{P}_{n-1}(x), \qquad n = 2, \ldots.$$

This, together with (1.54), leads to (1.51).

(3) This follows from Theorem 1.24.

(4) This follows from Theorem 1.22 and (1.50).

(5) The fact that $P_n(1) = 1$ follows from the three-term recurrence (1.51) 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 $n = 1$, (1.52) 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

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

Therefore, $((1-x^2)P_n'(x))'$ is orthogonal to all polynomials in $\mathcal{P}_{n-1}$. It thus follows from Corollary 1.23 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

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

These two equations imply (1.52) 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 $f_0(x) = 1$, $f_1(x) = 1$, and $f_2(x) = x^2$. Show that
   $$B_n f_0(x) = 1, \quad B_n f_1(x) = x, \quad B_n f_2(x) = \frac{n-1}{n} x^2 + \frac{1}{n} x, \qquad \forall x \in [0,1].$$
   (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 $\|(B_n f)' - f'\|_{C[a,b]} \to 0$ 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 $f, g \in C[a,b]$
   $$\|f + g\| \le \|f\| + \|g\| \quad \text{and} \quad \big|\|f\| - \|g\|\big| \le \|f - g\|.$$
5. Prove Theorem 1.3.
6. Let $n \ge 0$ be an integer and a, b two real numbers with $a < b$. Define
   $$F(c) = \max_{a \le x \le b} \left| \sum_{k=0}^{n} c_k x^k \right| \qquad \forall c = (c_0, \dots, c_n) \in \mathbb{R}^{n+1}.$$
   Show that $F : \mathbb{R}^{n+1} \to \mathbb{R}$ is a continuous function.
7. Let $k \ge 1$ be an integer and $f \in C^k[a,b]$. Show that, for any $\epsilon > 0$, there exists $p \in \mathcal{P}$ such that
   $$\|f - p\|_{C[a,b]} < \epsilon, \qquad \|f' - p'\|_{C[a,b]} < \epsilon, \qquad \dots, \qquad \|f^{(k)} - p^{(k)}\|_{C[a,b]} < \epsilon.$$
8. Let $f \in C[a,b]$ but $f \notin \mathcal{P}$. Show that there exits no polynomial $p \in \mathcal{P}$ such that
   $$\|f - p\|_{C[a,b]} \le \|f - q\|_{C[a,b]} \qquad \forall q \in \mathcal{P}.$$
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
   $$\max_{0 \le j \le n} |f(x_j) - p(x_j)| \le \max_{0 \le j \le n} |f(x_j) - q(x_j)| \qquad \forall q \in \mathcal{P}_m?$$
   Discuss the cases $0 \le m < n$, $m = n$, 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 $p \in \operatorname{Span}\{f_1, \dots, f_n\}$ such that
    $$\|f - p\|_{C[a,b]} \le \|f - q\|_{C[a,b]} \qquad \forall q \in \operatorname{Span}\{f_1, \dots, f_n\}.$$
    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

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

or such that

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

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

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

with $a_0, a_1, \ldots, a_n$ and $b_1, \ldots, b_n$ 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

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

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

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

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

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

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

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

25. Show that

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

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

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

be the Fourier coefficients of f. Let

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

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

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

28. Prove that

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

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 nth 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

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

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

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

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

(a) *Recursion formula.*

$$U_0(x) = 1, \quad U_1(x) = 2x,$$
$$U_{n+1}(x) = 2x U_n(x) - U_{n-1}(x), \qquad n = 1, \ldots;$$

(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.*

$$\left(1-x^2\right) U_n''(x) - 3x U_n'(x) + n(n+2) U_n(x) = 0 \qquad n = 0, \ldots;$$

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

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

(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

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

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

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

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

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

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

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

(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

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

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

$$\mu_n(f) = \int_a^b x^n f(x)\, dx, \qquad n = 0, \ldots.$$

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

40. Let

$$P_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} \left[\left(x^2 - 1 \right)^n \right], \qquad n = 0, \ldots$$

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.,

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

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 $\xi_1, \ldots, \xi_n$ 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 $\xi_1, \ldots, \xi_n$.

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

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

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.