4 Approximations

4.1 Background

In this chapter we are interested in approximation problems. Generally speaking, starting from a function $f(x)$ we would like to find a different function $g(x)$ that belongs to a given class of functions and is "close" to $f(x)$ in some sense. As far as the class of functions that $g(x)$ belongs to, we will typically assume that $g(x)$ is a polynomial of a given degree (though it can be a trigonometric function, or any other function). A typical approximation problem, will therefore be: find the "closest" polynomial of degree $\leqslant n$ to $f(x)$.

What do we mean by "close"? There are different ways of measuring the "distance" between two functions. We will focus on two such measurements (among many): the L^∞-norm and the L^2-norm. We chose to focus on these two examples because of the different mathematical techniques that are required to solve the corresponding approximation problems.

We start with several definitions. We recall that a **norm** on a vector space V over $\mathbb{R}$ is a function $\|\cdot\| : V \to \mathbb{R}$ with the following properties:

1. $\lambda\|f\| = |\lambda|\|f\|, \quad \forall\lambda \in \mathbb{R}$ and $\forall f \in V$.
2. $\|f\| \geqslant 0$, $\forall f \in V$. Also $\|f\| = 0$ iff f is the zero element of V.
3. The triangle inequality: $\|f + g\| \leqslant \|f\| + \|g\|$, $\forall f, g \in V$.

We assume that the function $f(x) \in C^0[a, b]$ (continuous on $[a, b]$). A continuous function on a closed interval obtains a maximum in the interval. We can therefore define **the L^∞ norm** (also known as **the maximum norm**) of such a function by

$$\|f\|_\infty = \max_{a\leqslant x\leqslant b} |f(x)|. \tag{4.1}$$

The L^∞-distance between two functions $f(x), g(x) \in C^0[a, b]$ is thus given by

$$\|f - g\|_\infty = \max_{a\leqslant x\leqslant b} |f(x) - g(x)|. \tag{4.2}$$

We note that the definition of the L^∞-norm can be extended to functions that are less regular than continuous functions. This generalization requires some subtleties that we would like to avoid in the following discussion, hence, we will limit ourselves to continuous functions.

We proceed by defining **the L^2-norm** of a continuous function $f(x)$ as

$$\|f\|_2 = \sqrt{\int_a^b |f(x)|^2 dx}. \tag{4.3}$$

The L^2 function space is the collection of functions $f(x)$ for which $\|f\|_2 < \infty$. Of course, we do not have to assume that $f(x)$ is continuous for the definition (4.3) to make sense. However, if we allow $f(x)$ to be discontinuous, we then have to be more rigorous in terms of the definition of the interval so that we end up with a norm (the problem is, e.g., in defining what is the "zero" element in the space). We therefore limit ourselves also in this case to continuous functions only. The L^2-distance between two functions $f(x)$ and $g(x)$ is

$$\|f-g\|_2 = \sqrt{\int_a^b |f(x)-g(x)|^2\,dx}. \tag{4.4}$$

At this point, a natural question is how important is the choice of norm in terms of the solution of the approximation problem. It is easy to see that the value of the norm of a function may vary substantially based on the function as well as the choice of the norm. For example, assume that $\|f\|_\infty < \infty$. Then, clearly

$$\|f\|_2 = \sqrt{\int_a^b |f|^2 dx} \leq (b-a)\|f\|_\infty.$$

On the other hand, it is easy to construct a function with an arbitrary small $\|f\|_2$ and an arbitrarily large $\|f\|_\infty$. Hence, the choice of norm may have a significant impact on the solution of the approximation problem.

As you have probably already anticipated, there is a strong connection between some approximation problems and interpolation problems. For example, one possible method of constructing an approximation to a given function is by sampling it at certain points and then interpolating the sampled data. Is that the best we can do? Sometimes the answer is positive, but the problem still remains difficult because we have to determine the best sampling points. We will address these issues in the following sections.

The following theorem, the Weierstrass approximation theorem, plays a central role in any discussion of approximations of functions. Loosely speaking, this theorem states that any continuous function can be approached as close as we want to with polynomials, assuming that the polynomials can be of any degree. We formulate this theorem in the L^∞ norm and note that a similar theorem holds also in the L^2 sense. We let Π_n denote the space of polynomials of degree $\leqslant n$.

Theorem 4.1 (Weierstrass Approximation Theorem) *Let $f(x)$ be a continuous function on $[a,b]$. Then there exists a sequence of polynomials $P_n(x)$ that converges uniformly to $f(x)$ on $[a,b]$, i.e., $\forall \varepsilon > 0$, there exists an $N \in \mathbb{N}$ and polynomials $P_n(x) \in \Pi_n$, such that $\forall x \in [a,b]$*

$$|f(x) - P_n(x)| < \varepsilon, \qquad \forall n \geqslant N.$$

We will provide a constructive proof of the Weierstrass approximation theorem: first, we will define a family of polynomials, known as **the Bernstein polynomials**, and then we will show that they uniformly converge to $f(x)$.

We start with the definition. Given a continuous function $f(x)$ in $[0,1]$, we define the Bernstein polynomials as

$$(B_n f)(x) = \sum_{j=0}^{n} f\left(\frac{j}{n}\right)\binom{n}{j} x^j (1-x)^{n-j}, \qquad 0 \leqslant x \leqslant 1.$$

We emphasize that the Bernstein polynomials depend on the function $f(x)$.

Example 4.2
Three Bernstein polynomials $B_6(x)$, $B_{10}(x)$, and $B_{20}(x)$ for the function

$$f(x) = \frac{1}{1+10(x-0.5)^2}$$

on the interval $[0,1]$ are shown in Figure 4.1. Note the gradual convergence of $B_n(x)$ to $f(x)$.

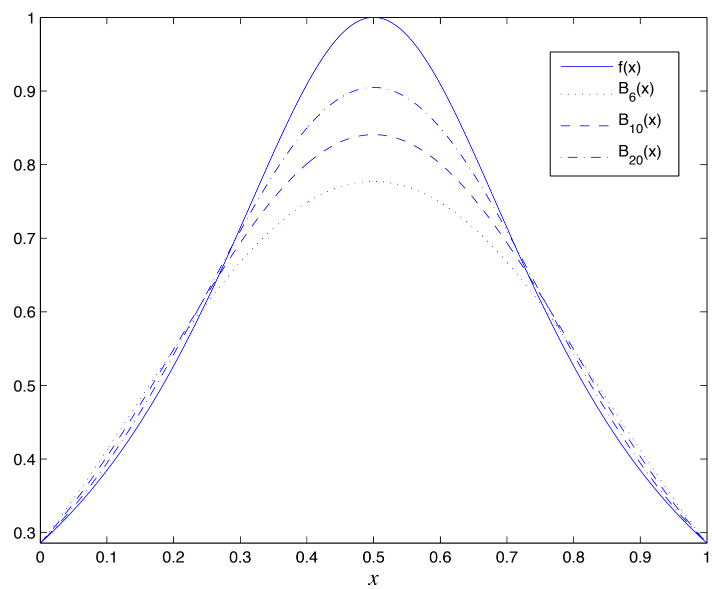

Figure 4.1: The Bernstein polynomials $B_6(x)$, $B_{10}(x)$, and $B_{20}(x)$ for the function $f(x) = \frac{1}{1+10(x-0.5)^2}$ on the interval $[0,1]$

We now state and prove several properties of $B_n(x)$ that will be used when we prove Theorem 4.1.

Lemma 4.3 *The following relations hold:*

1. $(B_n 1)(x) = 1$
2. $(B_n x)(x) = x$
3. $(B_n x^2)(x) = \frac{n-1}{n}x^2 + \frac{x}{n}$.

Proof.

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

$$\begin{aligned}
(B_n x)(x) &= \sum_{j=0}^{n} \frac{j}{n}\binom{n}{j} x^j (1-x)^{n-j} = x \sum_{j=1}^{n} \binom{n-1}{j-1} x^{j-1}(1-x)^{n-j} \\
&= x \sum_{j=0}^{n-1} \binom{n-1}{j} x^j (1-x)^{n-1-j} = x[x + (1-x)]^{n-1} = x.
\end{aligned}$$

Finally,

$$\begin{aligned}
\left(\frac{j}{n}\right)^2 \binom{n}{j} &= \frac{j}{n}\frac{(n-1)!}{(n-j)!(j-1)!} = \frac{n-1}{n-1}\frac{j-1}{n}\frac{(n-1)!}{(n-j)!(j-1)!} + \frac{1}{n}\frac{(n-1)!}{(n-j)!(j-1)!} \\
&= \frac{n-1}{n}\binom{n-2}{j-2} + \frac{1}{n}\binom{n-1}{j-1}.
\end{aligned}$$

Hence

$$\begin{aligned}
(B_n x^2)(x) &= \sum_{j=0}^{n} \left(\frac{j}{n}\right)^2 \binom{n}{j} x^j (1-x)^{n-j} \\
&= \frac{n-1}{n}x^2 \sum_{j=2}^{n} \binom{n-2}{j-2} x^{j-2}(1-x)^{n-j} + \frac{1}{n}x \sum_{j=1}^{n} \binom{n-1}{j-1} x^{j-1}(1-x)^{n-j} \\
&= \frac{n-1}{n}x^2 (x + (1-x))^{n-2} + \frac{1}{n}x(x + (1-x))^{n-1} = \frac{n-1}{n}x^2 + \frac{x}{n}. \quad \blacksquare
\end{aligned}$$

In the following lemma we state several additional properties of the Bernstein polynomials. The proof is left as an exercise.

Lemma 4.4 *For all functions $f(x), g(x)$ that are continuous in $[0,1]$, and $\forall \alpha \in \mathbb{R}$*

1. *Linearity.*

$$(B_n(\alpha f + g))(x) = \alpha(B_n f)(x) + (B_n g)(x).$$

2. *Monotonicity. If* $f(x) \leqslant g(x)$ $\forall x \in [0,1]$, *then*

$$(B_n f)(x) \leqslant (B_n g)(x).$$

Also, if $|f(x)| \leqslant g(x)$ $\forall x \in [0,1]$ *then*

$$|(B_n f)(x)| \leqslant (B_n g)(x).$$

3. *Positivity. If* $f(x) \geqslant 0$ *then*

$$(B_n f)(x) \geqslant 0.$$

We are now ready to prove the Weierstrass approximation theorem, Theorem 4.1.

Proof. We will prove the theorem in the interval $[0,1]$. The extension to $[a,b]$ is left as an exercise. Since $f(x)$ is continuous on a closed interval, it is uniformly continuous. Hence $\forall x, y \in [0,1]$, such that $|x-y| \leqslant \delta$,

$$|f(x) - f(y)| \leqslant \varepsilon. \tag{4.5}$$

In addition, since $f(x)$ is continuous on a closed interval, it is also bounded. Let

$$M = \max_{x \in [0,1]} |f(x)|.$$

Fix any point $a \in [0,1]$. If $|x-a| \leqslant \delta$ then (4.5) holds. If $|x-a| > \delta$ then

$$|f(x) - f(a)| \leqslant 2M \leqslant 2M \left(\frac{x-a}{\delta}\right)^2.$$

(at first sight this seems to be a strange way of upper bounding a function. We will use it later on to our advantage). Combining the estimates for both cases we have

$$|f(x) - f(a)| \leqslant \varepsilon + \frac{2M}{\delta^2}(x-a)^2.$$

We would now like to estimate the difference between $B_n f$ and f. The linearity of B_n and the property $(B_n 1)(x) = 1$ imply that

$$B_n(f - f(a))(x) = (B_n f)(x) - f(a).$$

Hence using the monotonicity of B_n and the mapping properties of x and x^2, we have,

$$\begin{aligned}
|B_n f(x) - f(a)| &\leqslant B_n\left(\varepsilon + \frac{2M}{\delta^2}(x-a)^2\right) = \varepsilon + \frac{2M}{\delta^2}\left(\frac{n-1}{n}x^2 + \frac{x}{n} - 2ax + a^2\right) \\
&= \varepsilon + \frac{2M}{\delta^2}(x-a)^2 + \frac{2M}{\delta^2}\frac{x-x^2}{n}.
\end{aligned}$$

Evaluating at $x = a$ we have (observing that $\max_{a\in[0,1]}(a - a^2) = \frac{1}{4}$)

$$|B_n f(a) - f(a)| \leqslant \varepsilon + \frac{2M}{\delta^2}\frac{a - a^2}{n} \leqslant \varepsilon + \frac{M}{2\delta^2 n}. \tag{4.6}$$

The point a was arbitrary so the result (4.6) holds for any point $a \in [0, 1]$. Choosing $N \geqslant \frac{M}{2\delta^2\varepsilon}$ we have $\forall n \geqslant N$,

$$\|B_n f - f\|_\infty \leqslant \varepsilon + \frac{M}{2\delta^2 N} \leqslant 2\varepsilon. \quad \blacksquare$$

- Is interpolation a good way of approximating functions in the ∞-norm? Not necessarily. Discuss Runge's example...

4.2 The Minimax Approximation Problem

We assume that the function $f(x)$ is continuous on $[a, b]$, and assume that $P_n(x)$ is a polynomial of degree $\leqslant n$. We recall that the L^∞-distance between $f(x)$ and $P_n(x)$ on the interval $[a, b]$ is given by

$$\|f - P_n\|_\infty = \max_{a\leqslant x\leqslant b} |f(x) - P_n(x)|. \tag{4.7}$$

Clearly, we can construct polynomials that will have an arbitrary large distance from $f(x)$. The question we would like to address is how close can we get to $f(x)$ (in the L^∞ sense) with polynomials of a given degree. We define $d_n(f)$ as the infimum of (4.7) over all polynomials of degree $\leqslant n$, i.e.,

$$d_n(f) = \inf_{P_n\in\Pi_n} \|f - P_n\|_\infty \tag{4.8}$$

The goal is to find a polynomial $P_n^*(x)$ for which the infimum (4.8) is actually obtained, i.e.,

$$d_n(f) = \|f - P_n^*(x)\|_\infty. \tag{4.9}$$

We will refer to a polynomial $P_n^*(x)$ that satisfies (4.9) as a **polynomial of best approximation** or **the minimax polynomial**. The minimal distance in (4.9) will be referred to as **the minimax error**.

The theory we will explore in the following sections will show that the minimax polynomial always exists and is unique. We will also provide a characterization of the minimax polynomial that will allow us to identify it if we actually see it. The general construction of the minimax polynomial will not be addressed in this text as it is relatively technically involved. We will limit ourselves to simple examples.

Example 4.5
We let $f(x)$ be a monotonically increasing and continuous function on the interval $[a, b]$ and are interested in finding the minimax polynomial of degree zero to $f(x)$ in that interval. We denote this minimax polynomial by

$$P_0^*(x) \equiv c.$$

Clearly, the smallest distance between $f(x)$ and P_0^* in the L^∞-norm will be obtained if

$$c = \frac{f(a) + f(b)}{2}.$$

The maximal distance between $f(x)$ and P_0^* will be attained at both edges and will be equal to

$$\pm \frac{f(b) - f(a)}{2}.$$

4.2.1 Existence of the minimax polynomial

The existence of the minimax polynomial is provided by the following theorem.

Theorem 4.6 (Existence) *Let $f \in C^0[a, b]$. Then for any $n \in \mathbb{N}$ there exists $P_n^*(x) \in \Pi_n$, that minimizes $\|f(x) - P_n(x)\|_\infty$ among all polynomials $P(x) \in \Pi_n$.*

Proof. We follow the proof as given in [**?**]. Let $\eta = (\eta_0, \ldots, \eta_n)$ be an arbitrary point in $\mathbb{R}^{n+1}$ and let

$$P_n(x) = \sum_{i=0}^{n} \eta_i x^i \in \Pi_n.$$

We also let

$$\phi(\eta) = \phi(\eta_0, \ldots, \eta_n) = \|f - P_n\|_\infty.$$

Our goal is to show that ϕ obtains a minimum in $\mathbb{R}^{n+1}$, i.e., that there exists a point $\eta^* = (\eta_0^*, \ldots, \eta_n^*)$ such that

$$\phi(\eta^*) = \min_{\eta \in \mathbb{R}^{n+1}} \phi(\eta).$$

Step 1. We first show that $\phi(\eta)$ is a continuous function on $\mathbb{R}^{n+1}$. For an arbitrary $\delta = (\delta_0, \ldots, \delta_n) \in \mathbb{R}^{n+1}$, define

$$q_n(x) = \sum_{i=0}^{n} \delta_i x^i.$$

Then

$$\phi(\eta+\delta) = \|f-(P_n+q_n)\|_\infty \leq \|f-P_n\|_\infty + \|q_n\|_\infty = \phi(\eta) + \|q_n\|_\infty.$$

Hence

$$\phi(\eta+\delta) - \phi(\eta) \leq \|q_n\|_\infty \leq \max_{x\in[a,b]}(|\delta_0| + |\delta_1||x| + \ldots + |\delta_n||x|^n).$$

For any $\varepsilon > 0$, let $\tilde{\delta} = \varepsilon/(1+c+\ldots+c^n)$, where $c = \max(|a|,|b|)$. Then for any $\delta = (\delta_0,\ldots,\delta_n)$ such that $\max|\delta_i| \leqslant \tilde{\delta}$, $0 \leqslant i \leqslant n$,

$$\phi(\eta+\delta) - \phi(\eta) \leqslant \varepsilon. \tag{4.10}$$

Similarly

$$\phi(\eta) = \|f-P_n\|_\infty = \|f-(P_n+q_n)+q_n\|_\infty \leqslant \|f-(P_n+q_n)\|_\infty + \|q_n\|_\infty = \phi(\eta+\delta) + \|q_n\|_\infty,$$

which implies that under the same conditions as in (4.10) we also get

$$\phi(\eta) - \phi(\eta+\delta) \leqslant \varepsilon,$$

Altogether,

$$|\phi(\eta+\delta) - \phi(\eta)| \leqslant \varepsilon,$$

which means that ϕ is continuous at η. Since η was an arbitrary point in $\mathbb{R}^{n+1}$, ϕ is continuous in the entire $\mathbb{R}^{n+1}$.

Step 2. We now construct a compact set in $\mathbb{R}^{n+1}$ on which ϕ obtains a minimum. We let

$$S = \left\{\eta \in \mathbb{R}^{n+1} \;\middle|\; \phi(\eta) \leq \|f\|_\infty\right\}.$$

We have

$$\phi(0) = \|f\|_\infty,$$

hence, $0 \in S$, and the set S is nonempty. We also note that the set S is bounded and closed (check!). Since ϕ is continuous on the entire $\mathbb{R}^{n+1}$, it is also continuous on S, and hence it must obtain a minimum on S, say at $\eta^* \in \mathbb{R}^{n+1}$, i.e.,

$$\min_{\eta\in S}\phi(\eta) = \phi(\eta^*).$$

Step 3. Since $0 \in S$, we know that

$$\min_{\eta \in S} \phi(\eta) \leqslant \phi(0) = \|f\|_\infty.$$

Hence, if $\eta \in \mathbb{R}^{n+1}$ but $\eta \notin S$ then

$$\phi(\eta) > \|f\|_\infty \geqslant \min_{\eta \in S} \phi(\eta).$$

This means that the minimum of ϕ over S is the same as the minimum over the entire $\mathbb{R}^{n+1}$. Therefore

$$P_n^*(x) = \sum_{i=0}^{n} \eta_i^* x^i, \tag{4.11}$$

is the best approximation of $f(x)$ in the L^∞ norm on $[a,b]$, i.e., it is the minimax polynomial, and hence the minimax polynomial exists. ■

We note that the proof of Theorem 4.6 is not a constructive proof. The proof does not tell us what the point η^* is, and hence, we do not know the coefficients of the minimax polynomial as written in (4.11). We will discuss the characterization of the minimax polynomial and some simple cases of its construction in the following sections.

4.2.2 Bounds on the minimax error

It is trivial to obtain an upper bound on the minimax error, since by the definition of $d_n(f)$ in (4.8) we have

$$d_n(f) \leqslant \|f - P_n\|_\infty, \qquad \forall P_n(x) \in \Pi_n.$$

A lower bound is provided by the following theorem.

Theorem 4.7 (de la Vallée-Poussin) *Let $a \leqslant x_0 < x_1 < \cdots < x_{n+1} \leqslant b$. Let $P_n(x)$ be a polynomial of degree $\leqslant n$. Suppose that*

$$f(x_j) - P_n(x_j) = (-1)^j e_j, \qquad j = 0, \ldots, n+1,$$

where all $e_j \neq 0$ and are of an identical sign. Then

$$\min_j |e_j| \leqslant d_n(f).$$

Proof. By contradiction. Assume for some $Q_n(x)$ that

$$\|f - Q_n\|_\infty < \min_j |e_j|.$$

Then the polynomial

$$(Q_n - P_n)(x) = (f - P_n)(x) - (f - Q_n)(x),$$

is a polynomial of degree $\leqslant n$ that has the same sign at x_j as does $f(x) - P_n(x)$. This implies that $(Q_n - P_n)(x)$ changes sign at least $n+2$ times, and hence it has at least $n+1$ zeros. Being a polynomial of degree $\leqslant n$ this is possible only if it is identically zero, i.e., if $P_n(x) \equiv Q_n(x)$, which contradicts the assumptions on $Q_n(x)$ and $P_n(x)$. ■

4.2.3 Characterization of the minimax polynomial

The following theorem provides a characterization of the minimax polynomial in terms of its oscillations property.

Theorem 4.8 (The oscillating theorem) *Suppose that $f(x)$ is continuous in $[a,b]$. The polynomial $P_n^*(x) \in \Pi_n$ is the minimax polynomial of degree n to $f(x)$ in $[a,b]$ if and only if $f(x) - P_n^*(x)$ assumes the values $\pm\|f - P_n^*\|_\infty$ with an alternating change of sign at least $n+2$ times in $[a,b]$.*

Proof. We prove here only the *sufficiency* part of the theorem. For the *necessary* part of the theorem we refer to [**?**].
Without loss of generality, suppose that

$$(f - P_n^*)(x_i) = (-1)^i \|f - P_n^*\|_\infty, \qquad 0 \leqslant i \leqslant n+1.$$

Let

$$D^* = \|f - P_n^*\|_\infty,$$

and let

$$d_n(f) = \min_{P_n \in \Pi_n} \|f - P_n\|_\infty.$$

We replace the infimum in the original definition of $d_n(f)$ by a minimum because we already know that a minimum exists. de la Vallée-Poussin's theorem (Theorem 4.7) implies that $D^* \leqslant d_n$. On the other hand, the definition of d_n implies that $d_n \leqslant D^*$. Hence $D^* = d_n$ and $P_n^*(x)$ is the minimax polynomial. ∎

Remark. In view of these theorems it is obvious why the Taylor expansion is a poor uniform approximation. The sum is non oscillatory.

4.2.4 Uniqueness of the minimax polynomial

Theorem 4.9 (Uniqueness) *Let $f(x)$ be continuous on $[a,b]$. Then its minimax polynomial $P_n^*(x) \in \Pi_n$ is unique.*

Proof. Let

$$d_n(f) = \min_{P_n \in \Pi_n} \|f - P_n\|_\infty.$$

Assume that $Q_n(x)$ is also a minimax polynomial. Then

$$\|f - P_n^*\|_\infty = \|f - Q_n\|_\infty = d_n(f).$$

The triangle inequality implies that

$$\|f - \frac{1}{2}(P_n^* + Q_n)\|_\infty \leq \frac{1}{2}\|f - P_n^*\|_\infty + \frac{1}{2}\|f - Q_n\|_\infty = d_n(f).$$

Hence, $\frac{1}{2}(P_n^* + Q_n) \in \Pi_n$ is also a minimax polynomial. The oscillating theorem (Theorem 4.8) implies that there exist $x_0, \ldots, x_{n+1} \in [a, b]$ such that

$$|f(x_i) - \frac{1}{2}(P_n^*(x_i) + Q_n(x_i))| = d_n(f), \qquad 0 \leqslant i \leqslant n+1. \tag{4.12}$$

Equation (4.12) can be rewritten as

$$|f(x_i) - P_n^*(x_i) + f(x_i) - Q_n(x_i)| = 2d_n(f), \qquad 0 \leqslant i \leqslant n+1. \tag{4.13}$$

Since $P_n^*(x)$ and $Q_n(x)$ are both minimax polynomials, we have

$$|f(x_i) - P_n^*(x_i)| \leq \|f - P_n^*\|_\infty = d_n(f), \qquad 0 \leqslant i \leqslant n+1. \tag{4.14}$$

and

$$|f(x_i) - Q_n(x_i)| \leq \|f - Q_n\|_\infty = d_n(f), \qquad 0 \leqslant i \leqslant n+1. \tag{4.15}$$

For any i, equations (4.13)–(4.15) mean that the absolute value of two numbers that are $\leqslant d_n(f)$ add up to $2d_n(f)$. This is possible only if they are equal to each other, i.e.,

$$f(x_i) - P_n^*(x_i) = f(x_i) - Q_n(x_i), \qquad 0 \leqslant i \leqslant n+1,$$

i.e.,

$$(P_n^* - Q_n)(x_i) = 0, \qquad 0 \leqslant i \leqslant n+1.$$

Hence, the polynomial $(P_n^* - Q_n)(x) \in \Pi_n$ has $n+2$ distinct roots which is possible for a polynomial of degree $\leqslant n$ only if it is identically zero. Hence

$$Q_n(x) \equiv P_n^*(x),$$

and the uniqueness of the minimax polynomial is established. ■

4.2.5 The near-minimax polynomial

We now connect between the minimax approximation problem and polynomial interpolation. In order for $f(x) - P_n(x)$ to change its sign $n+2$ times, there should be $n+1$ points on which $f(x)$ and $P_n(x)$ agree with each other. In other words, we can think of $P_n(x)$ as a function that interpolates $f(x)$ at (least in) $n+1$ points, say $x_0, \ldots, x_n$. What can we say about these points?

We recall that the interpolation error is given by (**??**),

$$f(x) - P_n(x) = \frac{1}{(n+1)!} f^{(n+1)}(\xi) \prod_{i=0}^{n} (x - x_i).$$

If $P_n(x)$ is indeed the minimax polynomial, we know that the maximum of

$$f^{(n+1)}(\xi) \prod_{i=0}^{n} (x - x_i), \tag{4.16}$$

will oscillate with equal values. Due to the dependency of $f^{(n+1)}(\xi)$ on the intermediate point ξ, we know that minimizing the error term (4.16) is a difficult task. We recall that interpolation at the Chebyshev points minimizes the multiplicative part of the error term, i.e.,

$$\prod_{i=0}^{n} (x - x_i).$$

Hence, choosing $x_0, \ldots, x_n$ to be the Chebyshev points will not result with the minimax polynomial, but nevertheless, this relation motivates us to refer to the interpolant at the Chebyshev points as the **near-minimax** polynomial. We note that the term "near-minimax" does not mean that the near-minimax polynomial is actually close to the minimax polynomial.

4.2.6 Construction of the minimax polynomial

The characterization of the minimax polynomial in terms of the number of points in which the maximum distance should be obtained with oscillating signs allows us to construct the minimax polynomial in simple cases by a direct computation.

We are not going to deal with the construction of the minimax polynomial in the general case. The algorithm for doing so is known as the Remez algorithm, and we refer the interested reader to [**?**] and the references therein.

A simple case where we can demonstrate a direct construction of the polynomial is when the function is convex, as done in the following example.

Example 4.10
Problem: Let $f(x) = e^x$, $x \in [1, 3]$. Find the minimax polynomial of degree $\leqslant 1$, $P_1^*(x)$.

Solution: Based on the characterization of the minimax polynomial, we will be looking for a linear function $P_1^*(x)$ such that its maximal distance between $P_1^*(x)$ and $f(x)$ is obtained 3 times with alternative signs. Clearly, in the case of the present problem, since the function is convex, the maximal distance will be obtained at both edges and at one interior point. We will use this observation in the construction that follows. The construction itself is graphically shown in Figure 4.2.

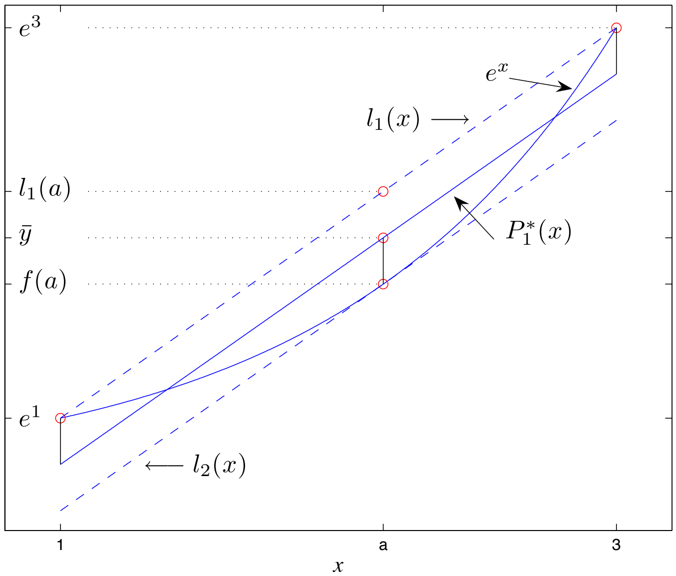

Figure 4.2: A construction of the linear minimax polynomial for the convex function e^x on $[1,3]$

We let $l_1(x)$ denote the line that connects the endpoints $(1,e)$ and $(3,e^3)$, i.e.,

$$l_1(x) = e + m(x-1).$$

Here, the slope m is given by

$$m = \frac{e^3 - e}{2}. \tag{4.17}$$

Let $l_2(x)$ denote the tangent to $f(x)$ at a point a that is identified such that the slope is m. Since $f'(x) = e^x$, we have $e^a = m$, i.e.,

$$a = \log m.$$

Now

$$f(a) = e^{\log m} = m,$$

and

$$l_1(a) = e + m(\log m - 1).$$

Hence, the average between $f(a)$ and $l_1(a)$ which we denote by $\bar{y}$ is given by

$$\bar{y} = \frac{f(a) + l_1(a)}{2} = \frac{m + e + m\log m - m}{2} = \frac{e + m\log m}{2}.$$

The minimax polynomial $P_1^*(x)$ is the line of slope m that passes through $(a, \bar{y})$,

$$P_1^*(x) - \frac{e + m\log m}{2} = m(x - \log m),$$

i.e.,

$$P_1^*(x) = mx + \frac{e - m\log m}{2},$$

where the slope m is given by (4.17). We note that the maximal difference between $P_1^*(x)$ and $f(x)$ is obtained at $x = 1, a, 3$.

4.3 Least-squares Approximations

4.3.1 The least-squares approximation problem

We recall that the L^2-norm of a function $f(x)$ is defined as

$$\|f\|_2 = \sqrt{\int_a^b |f(x)|^2 dx}.$$

As before, we let Π_n denote the space of all polynomials of degree $\leqslant n$. The **least-squares approximation problem** is to find the polynomial that is the closest to $f(x)$ in the L^2-norm among all polynomials of degree $\leqslant n$, i.e., to find $Q_n^* \in \Pi_n$ such that

$$\|f - Q_n^*\|_2 = \min_{Q_n \in \Pi_n} \|f - Q_n\|_2.$$

4.3.2 Solving the least-squares problem: a direct method

Our goal is to find a polynomial in Π_n that minimizes the distance $\|f(x) - Q_n(x)\|_2$ among all polynomials $Q_n \in \Pi_n$. We thus consider

$$Q_n(x) = \sum_{i=0}^{n} a_i x^i.$$

For convenience, instead of minimizing the L_2 norm of the difference, we will minimize its square. We thus let ϕ denote the square of the L_2-distance between $f(x)$ and $Q_n(x)$, i.e.,

$$\begin{aligned}
\phi(a_0, \ldots, a_n) &= \int_a^b (f(x) - Q_n(x))^2 dx \\
&= \int_a^b f^2(x)dx - 2\sum_{i=0}^{n} a_i \int_a^b x^i f(x)dx + \sum_{i=0}^{n}\sum_{j=0}^{n} a_i a_j \int_a^b x^{i+j} dx.
\end{aligned}$$

ϕ is a function of the $n+1$ coefficients in the polynomial $Q_n(x)$. This means that we want to find a point $a^* = (a_0^*, \ldots, a_n^*) \in \mathbb{R}^{n+1}$ for which ϕ obtains a minimum. At this point

$$\left.\frac{\partial \phi}{\partial a_k}\right|_{a=a^*} = 0. \tag{4.18}$$

The condition (4.18) implies that

$$\begin{aligned} 0 &= -2\int_a^b x^k f(x)dx + \sum_{i=0}^n a_i^* \int_a^b x^{i+k}dx + \sum_{j=0}^n a_j^* \int_a^b x^{j+k}dx \\ &= 2\left[\sum_{i=0}^n a_i^* \int_a^b x^{i+k}dx - \int_a^b x^k f(x)dx\right]. \end{aligned} \tag{4.19}$$

This is a linear system for the unknowns $(a_0^*, \ldots, a_n^*)$:

$$\sum_{i=0}^n a_i^* \int_a^b x^{i+k}dx = \int_a^b x^k f(x), \qquad k = 0, \ldots, n. \tag{4.20}$$

We thus know that the solution of the least-squares problem is the polynomial

$$Q_n^*(x) = \sum_{i=0}^n a_i^* x^i,$$

where the coefficients a_i^*, $i = 0, \ldots, n$, are the solution of (4.20), assuming that this system can be solved. Indeed, the system (4.20) always has a unique solution, which proves that not only the least-squares problem has a solution, but that it is also unique.

We let $H_{n+1}(a,b)$ denote the $(n+1)\times(n+1)$ coefficients matrix of the system (4.20) on the interval $[a,b]$, i.e.,

$$(H_{n+1}(a,b))_{i,k} = \int_a^b x^{i+k}dx, \qquad 0 \leqslant i,k \leqslant n.$$

For example, in the case where $[a,b] = [0,1]$,

$$H_n(0,1) = \begin{pmatrix} 1/1 & 1/2 & \ldots & 1/n \\ 1/2 & 1/3 & \ldots & 1/(n+1) \\ \vdots & & \vdots & \\ 1/n & 1/(n+1) & \ldots & 1/(2n-1) \end{pmatrix} \tag{4.21}$$

The matrix (4.21) is known as **the Hilbert matrix**.

Lemma 4.11 *The Hilbert matrix is invertible.*

Proof. We leave it is an exercise to show that the determinant of H_n is given by

$$\det(H_n) = \frac{(1!2!\cdots(n-1)!)^4}{1!2!\cdots(2n-1)!}.$$

Hence, $\det(H_n) \neq 0$ and H_n is invertible. ■

Is inverting the Hilbert matrix a good way of solving the least-squares problem? No. There are numerical instabilities that are associated with inverting H. We demonstrate this with the following example.

Example 4.12
The Hilbert matrix H_5 is

$$H_5 = \begin{pmatrix} 1/1 & 1/2 & 1/3 & 1/4 & 1/5 \\ 1/2 & 1/3 & 1/4 & 1/5 & 1/6 \\ 1/3 & 1/4 & 1/5 & 1/6 & 1/7 \\ 1/4 & 1/5 & 1/6 & 1/7 & 1/8 \\ 1/5 & 1/6 & 1/7 & 1/8 & 1/9 \end{pmatrix}$$

The inverse of H_5 is

$$H_5 = \begin{pmatrix} 25 & -300 & 1050 & -1400 & 630 \\ -300 & 4800 & -18900 & 26880 & -12600 \\ 1050 & -18900 & 79380 & -117600 & 56700 \\ -1400 & 26880 & -117600 & 179200 & -88200 \\ 630 & -12600 & 56700 & -88200 & 44100 \end{pmatrix}$$

The condition number of H_5 is $4.77 \cdot 10^5$, which indicates that it is ill-conditioned. In fact, the condition number of H_n increases with the dimension n so inverting it becomes more difficult with an increasing dimension.

4.3.3 Solving the least-squares problem: with orthogonal polynomials

Let $\{P_k\}_{k=0}^n$ be polynomials such that

$$\deg(P_k(x)) = k.$$

Let $Q_n(x)$ be a linear combination of the polynomials $\{P_k\}_{k=0}^n$, i.e.,

$$Q_n(x) = \sum_{j=0}^{n} c_j P_j(x). \tag{4.22}$$

Clearly, $Q_n(x)$ is a polynomial of degree $\leqslant n$. Define

$$\phi(c_0, \ldots, c_n) = \int_a^b [f(x) - Q_n(x)]^2 dx.$$

We note that the function ϕ is a quadratic function of the coefficients of the linear combination (4.22), $\{c_k\}$. We would like to minimize ϕ. Similarly to the calculations done in the previous section, at the minimum, $c^* = (c_0^*, \ldots, c_n^*)$, we have

$$0 = \left.\frac{\partial \phi}{\partial c_k}\right|_{c=c^*} = -2\int_a^b P_k(x)f(x)dx + 2\sum_{j=0}^{n} c_j^* \int_a^b P_j(x)P_k(x)dx,$$

i.e.,

$$\sum_{j=0}^{n} c_j^* \int_a^b P_j(x)P_k(x)dx = \int_a^b P_k(x)f(x)dx, \qquad k = 0, \ldots, n. \tag{4.23}$$

Note the similarity between equation (4.23) and (4.20). There, we used the basis functions $\{x^k\}_{k=0}^n$ (a basis of Π_n), while here we work with the polynomials $\{P_k(x)\}_{k=0}^n$ instead. The idea now is to choose the polynomials $\{P_k(x)\}_{k=0}^n$ such that the system (4.23) can be easily solved. This can be done if we choose them in such a way that

$$\int_a^b P_i(x)P_j(x)dx = \delta_{ij} = \begin{cases} 1, & i = j, \\ 0, & j \neq j. \end{cases} \tag{4.24}$$

Polynomials that satisfy (4.24) are called **orthonormal polynomials**. If, indeed, the polynomials $\{P_k(x)\}_{k=0}^n$ are orthonormal, then (4.23) implies that

$$c_j^* = \int_a^b P_j(x)f(x)dx, \qquad j = 0, \ldots, n. \tag{4.25}$$

The solution of the least-squares problem is a polynomial

$$Q_n^*(x) = \sum_{j=0}^{n} c_j^* P_j(x), \tag{4.26}$$

with coefficients c_j^*, $j = 0, \ldots, n$, that are given by (4.25).

Remark. Polynomials that satisfy

$$\int_a^b P_i(x)P_j(x)dx = \begin{cases} \int_a^b (P_i(x))^2, & i = j, \\ \\ 0, & i \neq j, \end{cases}$$

with $\int_a^b (P_i(x))^2 dx$ that is not necessarily 1 are called **orthogonal polynomials**. In this case, the solution of the least-squares problem is given by the polynomial $Q_n^*(x)$ in (4.26) with the coefficients

$$c_j^* = \frac{\int_a^b P_j(x)f(x)dx}{\int_a^b (P_j(x))^2 dx}, \qquad j = 0, \ldots, n. \tag{4.27}$$

4.3.4 The weighted least squares problem

A more general least-squares problem is the **weighted least squares approximation problem**. We consider a **weight function**, $w(x)$, to be a continuous on (a, b), non-negative function with a positive mass, i.e.,

$$\int_a^b w(x)dx > 0.$$

Note that $w(x)$ may be singular at the edges of the interval since we do not require it to be continuous on the closed interval $[a, b]$. For any weight $w(x)$, we define the corresponding weighted L^2-norm of a function $f(x)$ as

$$\|f\|_{2,w} = \sqrt{\int_a^b (f(x))^2 w(x)dx}.$$

The weighted least-squares problem is to find the closest polynomial $Q_n^* \in \Pi_n$ to $f(x)$, this time in the weighted L^2-norm sense, i.e., we look for a polynomial $Q_n^*(x)$ of degree $\leqslant n$ such that

$$\|f - Q_n^*\|_{2,w} = \min_{Q_n \in \Pi_n} \|f - Q_n\|_{2,w}. \tag{4.28}$$

In order to solve the weighted least-squares problem (4.28) we follow the methodology described in Section 4.3.3, and consider polynomials $\{P_k\}_{k=0}^n$ such that $\deg(P_k(x)) = k$. We then consider a polynomial $Q_n(x)$ that is written as their linear combination:

$$Q_n(x) = \sum_{j=0}^n c_j P_j(x).$$

By repeating the calculations of Section 4.3.3, we obtain for the coefficients of the minimizer Q_n^*,

$$\sum_{j=0}^n c_j^* \int_a^b w(x)P_j(x)P_k(x)dx = \int_a^b w(x)P_k(x)f(x)dx, \qquad k = 0, \ldots, n, \tag{4.29}$$

(compare with (4.23)). The system (4.29) can be easily solved if we choose $\{P_k(x)\}$ to be orthonormal with respect to the weight $w(x)$, i.e.,

$$\int_a^b P_i(x)P_j(x)w(x)dx = \delta_{ij}.$$

Hence, the solution of the weighted least-squares problem is given by

$$Q_n^*(x) = \sum_{j=0}^n c_j^* P_j(x), \tag{4.30}$$

where the coefficients are given by

$$c_j^* = \int_a^b P_j(x) f(x) w(x) dx, \qquad j = 0, \ldots, n. \tag{4.31}$$

Remark. In the case where $\{P_k(x)\}$ are orthogonal but not necessarily normalized, the solution of the weighted least-squares problem is given by

$$Q_n^*(x) = \sum_{j=0}^{n} c_j^* P_j(x)$$

with

$$c_j^* = \frac{\int_a^b P_j(x) f(x) w(x) dx}{\int_a^b (P_j(x))^2 w(x) dx}, \qquad j = 0, \ldots, n.$$

4.3.5 Orthogonal polynomials

At this point we already know that orthogonal polynomials play a central role in the solution of least-squares problems. In this section we will focus on the construction of orthogonal polynomials. The properties of orthogonal polynomials will be studies in Section 4.4.2.

We start by defining the **weighted inner product** between two functions $f(x)$ and $g(x)$ (with respect to the weight $w(x)$):

$$\langle f, g \rangle_w = \int_a^b f(x) g(x) w(x) dx.$$

To simplify the notations, even in the weighted case, we will typically write $\langle f, g \rangle$ instead of $\langle f, g \rangle_w$. Some properties of the weighted inner product include

1. $\langle \alpha f, g \rangle = \langle f, \alpha g \rangle = \alpha \langle f, g \rangle, \quad \forall \alpha \in \mathbb{R}.$
2. $\langle f_1 + f_2, g \rangle = \langle f_1, g \rangle + \langle f_2, g \rangle.$
3. $\langle f, g \rangle = \langle g, f \rangle$
4. $\langle f, f \rangle \geqslant 0$ and $\langle f, f \rangle = 0$ iff $f \equiv 0$. Here we must assume that $f(x)$ is continuous in the interval $[a, b]$. If it is not continuous, we can have $\langle f, f \rangle = 0$ and $f(x)$ can still be non-zero (e.g., in one point).

The weighted L_2-norm can be obtained from the weighted inner product by

$$\|f\|_{2,w} = \sqrt{\langle f, f \rangle_w}.$$

Given a weight $w(x)$, we are interested in constructing orthogonal (or orthonormal) polynomials. This can be done using **the Gram-Schmidt orthogonalization process**, which we now describe in detail.

In the general context of linear algebra, the Gram-Schmidt process is being used to convert one set of linearly independent vectors to an orthogonal set of vectors that spans the same subspace as the original set. In our context, we should think about the process as converting one set of polynomials that span the space of polynomials of degree $\leqslant n$ to an orthogonal set of polynomials that spans the same space Π_n. Accordingly, we set the initial set of polynomials as $\{1, x, x^2, \ldots, x^n\}$, which we would like to convert to orthogonal polynomials (of an increasing degree) with respect to the weight $w(x)$.

We will first demonstrate the process with the weight $w(x) \equiv 1$. We will generate a set of orthogonal polynomials $\{P_0(x), \ldots, P_n(x)\}$ from $\{1, x, \ldots, x^n\}$. The degree of the polynomials P_i is i.

We start by setting

$$P_0(x) = 1.$$

We then let

$$P_1(x) = x - c_1^0 P_0(x) = x - c_1^0.$$

The orthogonality condition $\int_a^b P_1 P_0 dx = 0$, implies that

$$\int_a^b 1 \cdot (x - c_1^0) dx = 0,$$

from which c= (a+b)/2, and thus

$$P_1(x) = x - \frac{a+b}{2}.$$

The computation continues in a similar fashion. We set

$$P_2(x) = x^2 - c_2^0 P_0(x) - c_2^1(x).$$

The two unknown coefficients, c_2^0 and c_2^1, are computed from the orthogonality conditions. This time, $P_2(x)$ should be orthogonal to $P_0(x)$ and to $P_1(x)$, i.e.,

$$\int_a^b P_2(x) P_0(x) dx = 0, \qquad \text{and} \qquad \int_a^b P_2(x) P_1(x) dx = 0,$$

and so on. If, in addition to the orthogonality condition, we would like the polynomials to be orthonormal, all that remains is to normalize:

$$\hat{P}_n(x) = \frac{P_n(x)}{\|P_n(x)\|} = \frac{P_n(x)}{\sqrt{\int_a^b (P_n(x))^2 dx}}, \quad \forall n.$$

The orthogonalization process is identical to the process that we described even when the weight $w(x)$ is not uniformly one. In this case, every integral will contain the weight.

4.4 Examples of orthogonal polynomials

This section includes several examples of orthogonal polynomials and a very brief summary of some of their properties.

1. **Legendre polynomials.** We start with the Legendre polynomials. This is a family of polynomials that are orthogonal with respect to the weight

   $$w(x) \equiv 1,$$

   on the interval $[-1, 1]$.

   In addition to deriving the Legendre polynomials through the Gram-Schmidt orthogonalization process, it can be shown that the Legendre polynomials can be obtained from the recurrence relation

   $$(n+1)P_{n+1}(x) - (2n+1)xP_n(x) + nP_{n-1}(x) = 0, \qquad n \geqslant 1, \tag{4.32}$$

   starting from the first two polynomials:

   $$P_0(x) = 1, \qquad P_1(x) = x.$$

   Instead of calculating these polynomials one by one from the recurrence relation, they can be obtained directly from Rodrigues' formula

   $$P_n(x) = \frac{1}{2^n n!}\frac{d^n}{dx^n}\left[(x^2-1)^n\right], \quad n \geqslant 0. \tag{4.33}$$

   The Legendre polynomials satisfy the orthogonality condition

   $$\langle P_n, P_m \rangle = \frac{2}{2n+1}\delta_{nm}. \tag{4.34}$$

2. **Chebyshev polynomials.** Our second example is of the Chebyshev polynomials. These polynomials are orthogonal with respect to the weight

   $$w(x) = \frac{1}{\sqrt{1-x^2}},$$

   on the interval $[-1, 1]$. They satisfy the recurrence relation

   $$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x), \quad n \geqslant 1, \tag{4.35}$$

   together with $T_0(x) = 1$ and $T_1(x) = x$ (see (**??**)). They and are explicitly given by

   $$T_n(x) = \cos(n \cos^{-1} x), \qquad n \geqslant 0. \tag{4.36}$$

(see (**??**)). The orthogonality relation that they satisfy is

$$\langle T_n, T_m \rangle = \begin{cases} 0, & n \neq m, \\ \pi, & n = m = 0, \\ \frac{\pi}{2}, & n = m \neq 0. \end{cases} \tag{4.37}$$

3. **Laguerre polynomials.** We proceed with the Laguerre polynomials. Here the interval is given by $[0, \infty)$ with the weight function

$$w(x) = e^{-x}.$$

The Laguerre polynomials are given by

$$L_n(x) = \frac{e^x}{n!} \frac{d^n}{dx^n}(x^n e^{-x}), \quad n \geqslant 0. \tag{4.38}$$

The normalization condition is

$$\|L_n\| = 1. \tag{4.39}$$

A more general form of the Laguerre polynomials is obtained when the weight is taken as

$$e^{-x} x^{\alpha},$$

for an arbitrary real $\alpha > -1$, on the interval $[0, \infty)$.

4. **Hermite polynomials.** The Hermite polynomials are orthogonal with respect to the weight

$$w(x) = e^{-x^2},$$

on the interval $(-\infty, \infty)$. The can be explicitly written as

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n e^{-x^2}}{dx^n}, \qquad n \geqslant 0. \tag{4.40}$$

Another way of expressing them is by

$$H_n(x) = \sum_{k=0}^{[n/2]} \frac{(-1)^k n!}{k!(n-2k)!} (2x)^{n-2k}, \tag{4.41}$$

where $[x]$ denotes the largest integer that is $\leqslant x$. The Hermite polynomials satisfy the recurrence relation

$$H_{n+1}(x) - 2xH_n(x) + 2nH_{n-1}(x) = 0, \qquad n \geqslant 1, \tag{4.42}$$

together with

$$H_0(x) = 1, \qquad H_1(x) = 2x.$$

They satisfy the orthogonality relation

$$\int_{-\infty}^{\infty} e^{-x^2} H_n(x) H_m(x) dx = 2^n n! \sqrt{\pi} \delta_{nm}. \tag{4.43}$$

4.4.1 Another approach to the least-squares problem

In this section we present yet another way of deriving the solution of the least-squares problem. Along the way, we will be able to derive some new results. We recall that our goal is to minimize $\|f(x) - Q_n(x)\|_{2,w}$, $\forall Q_n \in \Pi_n$, i.e., to minimize the integral

$$\int_a^b w(x)(f(x) - Q_n(x))^2 dx \tag{4.44}$$

among all the polynomials $Q_n(x)$ of degree $\leqslant n$. The minimizer of (4.44) is denoted by $Q_n^*(x)$.

Assume that $\{P_k(x)\}_{k\geqslant 0}$ is an orthogonal family of polynomials with respect to $w(x)$, and let

$$Q_n(x) = \sum_{j=0}^{n} c_j P_j(x).$$

Then

$$\|f - Q_n\|_{2,w}^2 = \int_a^b w(x) \left(f(x) - \sum_{j=0}^{n} c_j P_j(x) \right)^2 dx.$$

Hence

$$\begin{aligned} 0 \;\leqslant\; & \left\langle f - \sum_{j=0}^{n} c_j P_j, f - \sum_{j=0}^{n} c_j P_j \right\rangle_w = \langle f, f \rangle_w - 2 \sum_{j=0}^{n} c_j \langle f, P_j \rangle_w + \sum_{i=0}^{n} \sum_{j=0}^{n} c_i c_j \langle P_i, P_j \rangle_w \\ & = \|f\|_{2,w}^2 - 2 \sum_{j=0}^{n} c_j \langle f, P_j \rangle_w + \sum_{j=0}^{n} c_j^2 \|P_j\|_{2,w}^2 \\ & = \|f\|_{2,w}^2 - \sum_{j=0}^{n} \frac{\langle f, P_j \rangle_w^2}{\|P_j\|_{2,w}^2} + \sum_{j=0}^{n} \left(\frac{\langle f, P_j \rangle_w}{\|P_j\|_{2,w}} - c_j \|P_j\|_{2,w} \right)^2. \end{aligned}$$

The last expression is minimal iff

$$\frac{\langle f, P_j \rangle_w}{\|P_j\|_{2,w}} - c_j \|P_j\|_{2,w} = 0, \qquad \forall 0 \leqslant j \leqslant n,$$

i.e., if

$$c_j = \frac{\langle f, P_j \rangle_w}{\|P_j\|_{2,w}^2}.$$

Hence, there exists a unique least-squares approximation which is given by

$$Q_n^*(x) = \sum_{j=0}^{n} \frac{\langle f, P_j \rangle_w}{\|P_j\|_{2,w}^2} P_j(x). \tag{4.45}$$

If the polynomials $\{P_j(x)\}$ are also normalized so that $\|P_j\|_{2,w} = 1$, then the minimizer $Q_n^*(x)$ in (4.45) becomes

$$Q_n^*(x) = \sum_{j=0}^{n} \langle f, P_j \rangle_w P_j(x).$$

Remarks.

1. We can write

$$\begin{aligned}\|f - Q_n^*\|_{2,w}^2 &= \int_a^b w(x) \left(f(x) - \sum_{j=0}^{n} c_j P_j(x) \right)^2 dx = \\ &= \|f\|_{2,w}^2 - 2 \sum_{j=0}^{n} \langle f, P_j \rangle_w c_j + \sum_{j=0}^{n} \|P_j\|_{2,w}^2 c_j^2.\end{aligned}$$

Since $\|P_j\|_{2,w} = 1$, $c_j = \langle f, P_j \rangle_w$, so that

$$\|f - Q_n^*\|_{2,w}^2 = \|f\|_{2,w}^2 - \sum_{j=0}^{n} \langle f, P_j \rangle_w^2 .$$

Hence

$$\sum_{j=0}^{n} \langle f, P_j \rangle_w^2 = \|f\|^2 - \|f - Q_n^*\|^2 \leqslant \|f\|^2,$$

i.e.,

$$\sum_{j=0}^{n} \langle f, P_j \rangle_w^2 \leqslant \|f\|_{2,w}^2. \tag{4.46}$$

The inequality (4.46) is called **Bessel's inequality**.

2. Assuming that $[a, b]$ is finite, we have

$$\lim_{n\to\infty} \|f - Q_n^*\|_{2,w} = 0.$$

Hence

$$\|f\|_{2,w}^2 = \sum_{j=0}^{\infty} \langle f, P_j \rangle_w^2, \tag{4.47}$$

which is known as **Parseval's equality**.

Example 4.13

Problem: Let $f(x) = \cos x$ on $[-1, 1]$. Find the polynomial in Π_2, that minimizes

$$\int_{-1}^{1} [f(x) - Q_2(x)]^2 dx.$$

Solution: The weight $w(x) \equiv 1$ on $[-1, 1]$ implies that the orthogonal polynomials we need to use are the Legendre polynomials. We are seeking for polynomials of degree $\leqslant 2$ so we write the first three Legendre polynomials

$$P_0(x) \equiv 1, \quad P_1(x) = x, \quad P_2(x) = \frac{1}{2}(3x^2 - 1).$$

The normalization factor satisfies, in general,

$$\int_{-1}^{1} P_n^2(x) = \frac{2}{2n+1}.$$

Hence

$$\int_{-1}^{1} P_0^2(x)dx = 2, \quad \int_{-1}^{1} P_1(x)dx = \frac{2}{3}, \quad \int_{-1}^{1} P_2^2(x)dx = \frac{2}{5}.$$

We can then replace the Legendre polynomials by their normalized counterparts:

$$P_0(x) \equiv \frac{1}{\sqrt{2}}, \quad P_1(x) = \sqrt{\frac{3}{2}}x, \quad P_2(x) = \frac{\sqrt{5}}{2\sqrt{2}}(3x^2 - 1).$$

We now have

$$\langle f, P_0 \rangle = \int_{-1}^{1} \cos x \frac{1}{\sqrt{2}} dx = \frac{1}{\sqrt{2}} \sin x \Big|_{-1}^{1} = \sqrt{2} \sin 1.$$

Hence

$$Q_0^*(x) \equiv \sin 1.$$

We also have

$$\langle f, P_1 \rangle = \int_{-1}^{1} \cos x \sqrt{\frac{3}{2}} x dx = 0.$$

which means that $Q_1^*(x) = Q_0^*(x)$. Finally,

$$\langle f, P_2 \rangle = \int_{-1}^{1} \cos x \sqrt{\frac{5}{2}} \frac{3x^2 - 1}{2} = \frac{1}{2}\sqrt{\frac{5}{2}}(12\cos 1 - 8\sin 1),$$

and hence the desired polynomial, $Q_2^*(x)$, is given by

$$Q_2^*(x) = \sin 1 + \left(\frac{15}{2}\cos 1 - 5\sin 1\right)(3x^2 - 1).$$

In Figure 4.3 we plot the original function $f(x) = \cos x$ (solid line) and its approximation $Q_2^*(x)$ (dashed line). We zoom on the interval $x \in [0, 1]$.

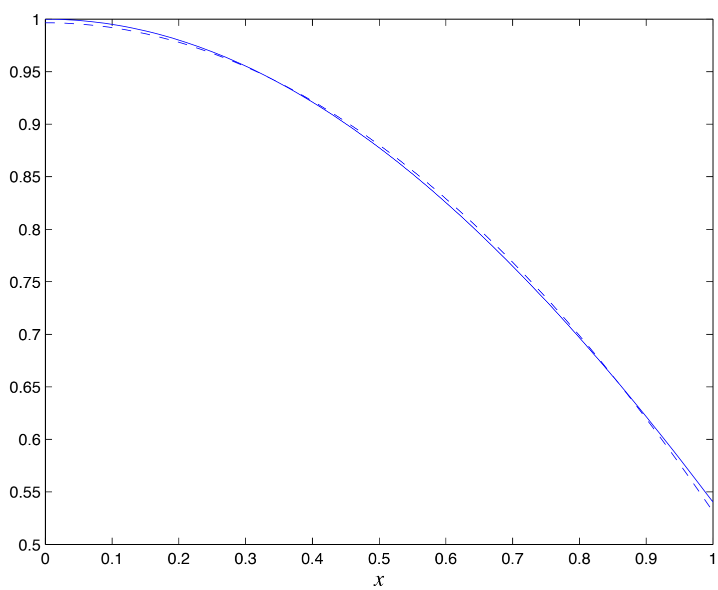

Figure 4.3: A second-order L^2-approximation of $f(x) = \cos x$. *Solid line:* $f(x)$; *Dashed line:* its approximation $Q_2^*(x)$

If the weight is $w(x) \equiv 1$ but the interval is $[a, b]$, we can still use the Legendre polynomials if we make the following change of variables. Define

$$x = \frac{b + a + (b - a)t}{2}.$$

Then the interval $-1 \leqslant t \leqslant 1$ is mapped to $a \leqslant x \leqslant b$. Now, define

$$F(t) = f\left(\frac{b+a+(b-a)t}{2}\right) = f(x).$$

Hence

$$\int_a^b [f(x) - Q_n(x)]^2 dx = \frac{b-a}{2}\int_{-1}^1 [F(t) - q_n(t)]^2 dt.$$

Example 4.14
Problem: Let $f(x) = \cos x$ on $[0, \pi]$. Find the polynomial in Π_1 that minimizes

$$\int_0^\pi [f(x) - Q_1(x)]^2 dx.$$

Solution:

$$\int_0^\pi (f(x) - Q_1^*(x))^2 dx = \frac{\pi}{2}\int_{-1}^1 [F(t) - q_n(t)]^2 dt.$$

Letting

$$x = \frac{\pi + \pi t}{2} = \frac{\pi}{2}(1+t),$$

we have

$$F(t) = \cos\left(\frac{\pi}{2}(1+t)\right) = -\sin\frac{\pi t}{2}.$$

We already know that the first two normalized Legendre polynomials are

$$P_0(t) = \frac{1}{\sqrt{2}}, \quad P_1(t) = \sqrt{\frac{3}{2}}t.$$

Hence

$$\langle F, P_0\rangle = -\int_{-1}^1 \frac{1}{\sqrt{2}}\sin\frac{\pi t}{2} dt = 0,$$

which means that $Q_0^*(t) = 0$. Also

$$\langle F, P_1\rangle = -\int_{-1}^1 \sin\frac{\pi t}{2}\sqrt{\frac{3}{2}} t dt = -\sqrt{\frac{3}{2}}\left[\frac{\sin\frac{\pi t}{2}}{\left(\frac{\pi}{2}\right)^2} - \frac{t\cos\frac{\pi t}{2}}{\frac{\pi}{2}}\right]_{-1}^1 = -\sqrt{\frac{3}{2}}\cdot\frac{8}{\pi^2}.$$

Hence

$$q_1^*(t) = -\frac{3}{2}\cdot\frac{8}{\pi^2}t = -\frac{12}{\pi^2}t \quad \Longrightarrow \quad Q_1^*(x) = -\frac{12}{\pi^2}\left(\frac{2}{\pi}x - 1\right).$$

In Figure 4.4 we plot the original function $f(x) = \cos x$ (solid line) and its approximation $Q_1^*(x)$ (dashed line).

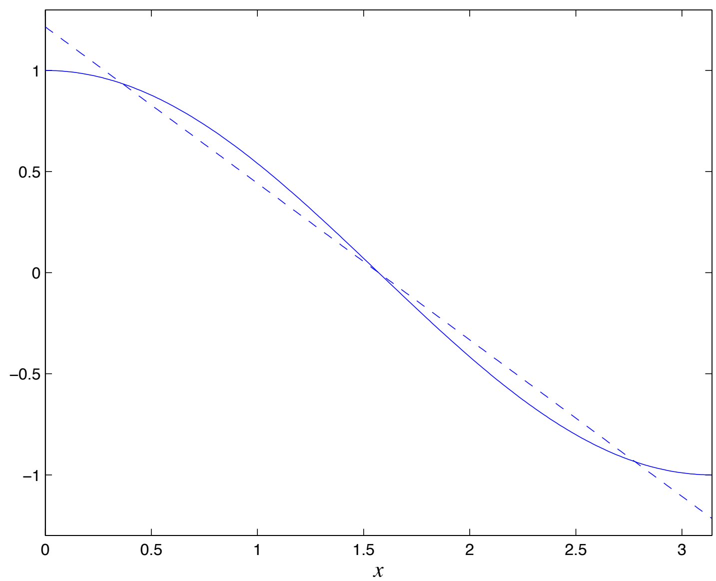

Figure 4.4: A first-order L^2-approximation of $f(x) = \cos x$ on the interval $[0, \pi]$. *Solid line:* $f(x)$, *Dashed line:* its approximation $Q_1^*(x)$

Example 4.15
Problem: Let $f(x) = \cos x$ in $[0, \infty)$. Find the polynomial in Π_1 that minimizes

$$\int_0^\infty e^{-x}[f(x) - Q_1(x)]^2 dx.$$

Solution: The family of orthogonal polynomials that correspond to this weight on $[0, \infty)$ are Laguerre polynomials. Since we are looking for the minimizer of the weighted L_2 norm among polynomials of degree $\leqslant 1$, we will need to use the first two Laguerre polynomials:

$$L_0(x) = 1, \qquad L_1(x) = 1 - x.$$

We thus have

$$\langle f, L_0 \rangle_w = \int_0^\infty e^{-x} \cos x dx = \left. \frac{e^{-x}(-\cos x + \sin x)}{2} \right|_0^\infty = \frac{1}{2}.$$

Also

$$\langle f, L_1 \rangle_w = \int_0^\infty e^{-x} \cos x (1-x) dx = \frac{1}{2} - \left[\frac{xe^{-x}(-\cos x + \sin x)}{2} - \frac{e^{-x}(-2\sin x)}{4} \right]_0^\infty = 0.$$

This means that

$$Q_1^*(x) = \langle f, L_0 \rangle_w L_0(x) + \langle f, L_1 \rangle_w L_1(x) = \frac{1}{2}.$$

4.4.2 Properties of orthogonal polynomials

We start with a theorem that deals with some of the properties of the roots of orthogonal polynomials. This theorem will become handy when we discuss Gaussian quadratures in Section **??**. We let $\{P_n(x)\}_{n\geqslant 0}$ be orthogonal polynomials in $[a, b]$ with respect to the weight $w(x)$.

Theorem 4.16 *The roots x_j, $j = 1, \ldots, n$ of $P_n(x)$ are all real, simple, and are in (a, b).*

Proof. Let $x_1, \ldots, x_r$ be the roots of $P_n(x)$ in (a, b). Let

$$Q_r(x) = (x - x_1) \cdot \ldots \cdot (x - x_r).$$

Then $Q_r(x)$ and $P_n(x)$ change their signs together in (a, b). Also

$$\deg(Q_r(x)) = r \leqslant n.$$

Hence $(P_n Q_r)(x)$ is a polynomial with one sign in (a, b). This implies that

$$\int_a^b P_n(x)Q_r(x)w(x)dx \neq 0,$$

and hence $r = n$ since $P_n(x)$ is orthogonal to polynomials of degree less than n. Without loss of generality we now assume that x_1 is a multiple root, i.e.,

$$P_n(x) = (x - x_1)^2 P_{n-2}(x).$$

Hence

$$P_n(x)P_{n-2}(x) = \left(\frac{P_n(x)}{x - x_1}\right)^2 \geqslant 0,$$

which implies that

$$\int_a^b P_n(x)P_{n-2}(x)dx > 0.$$

This is not possible since P_n is orthogonal to P_{n-2}. Hence roots can not repeat. ∎

Another important property of orthogonal polynomials is that they can all be written in terms of recursion relations. We have already seen specific examples of such relations for the Legendre, Chebyshev, and Hermite polynomials (see (4.32), (4.35), and (4.42)). The following theorem states such relations always hold.

Theorem 4.17 (Triple Recursion Relation) *Any three consecutive orthonormal polynomials are related by a recursion formula of the form*

$$P_{n+1}(x) = (A_n x + B_n)P_n(x) - C_n P_{n-1}(x).$$

If a_k and b_k are the coefficients of the terms of degree k and degree $k-1$ in $P_k(x)$, then

$$A_n = \frac{a_{n+1}}{a_n}, \quad B_n = \frac{a_{n+1}}{a_n}\left(\frac{b_{n+1}}{a_{n+1}} - \frac{b_n}{a_n}\right), \quad C_n = \frac{a_{n+1}a_{n-1}}{a_n^2}.$$

Proof. For

$$A_n = \frac{a_{n+1}}{a_n},$$

let

$$\begin{aligned} Q_n(x) &= P_{n+1}(x) - A_n x P_n(x) \\ &= (a_{n+1}x^{n+1} + b_{n+1}x^n + \ldots) - \frac{a_{n+1}}{a_n} x(a_n x^n + b_n x^{n-1} + \ldots) \\ &= \left(b_{n+1} - \frac{a_{n+1}b_n}{a_n}\right) x^n + \ldots \end{aligned}$$

Hence $\deg(Q(x)) \leqslant n$, which means that there exists $\alpha_0, \ldots, \alpha_n$ such that

$$Q(x) = \sum_{i=0}^{n} \alpha_i P_i(x).$$

For $0 \leqslant i \leqslant n-2$,

$$\alpha_i = \frac{\langle Q, P_i\rangle}{\langle P_i, P_i\rangle} = \langle Q, P_i\rangle = \langle P_{n+1} - A_n x P_n, P_i\rangle = -A_n \langle xP_n, P_i\rangle = 0.$$

Hence

$$Q_n(x) = \alpha_n P_n(x) + \alpha_{n-1}P_{n-1}(x).$$

Set $\alpha_n = B_n$ and $\alpha_{n-1} = -C_n$. Then, since

$$xP_{n-1} = \frac{a_{n-1}}{a_n} P_n + q_{n-1},$$

we have

$$C_n = A_n \langle xP_n, P_{n-1}\rangle = A_n \langle P_n, xP_{n-1}\rangle = A_n \left\langle P_n, \frac{a_{n-1}}{a_n}P_n + q_{n-1}\right\rangle = A_n \frac{a_{n-1}}{a_n}.$$

Finally

$$P_{n+1} = (A_n x + B_n)P_n - C_n P_{n-1},$$

can be explicitly written as

$$a_{n+1}x^{n+1} + b_{n+1}x^n + \ldots = (A_n x + B_n)(a_n x^n + b_n x^{n-1} + \ldots) - C_n(a_{n-1}x^{n-1} + b_{n-1}x^{n-2} + \ldots).$$

The coefficient of x^n is

$$b_{n+1} = A_n b_n + B_n a_n,$$

which means that

$$B_n = (b_{n+1} - A_n b_n)\frac{1}{a_n}. \quad \blacksquare$$