8

Polynomial approximation in the ∞-norm

8.1 Introduction

In Chapter 6 we considered the problem of interpolating a function by polynomials of a certain degree. Here we shall discuss other types of approximation by polynomials, the overall objective being to find the polynomial of given degree n which provides the 'best approximation' from $\mathcal{P}_n$ to a given function in a sense that will be made precise below.

8.2 Normed linear spaces

In order to be able to talk about 'best approximation' in a rigorous manner we need to recall from Chapter 2 the concept of *norm*; this will allow us to compare various approximations quantitatively and select the one which has the smallest approximation error. The definition given in Section 2.7 applies to a linear space consisting of functions in the same way as to the finite-dimensional linear spaces considered in Chapter 2.

Definition 8.1 *Suppose that $\mathcal{V}$ is a linear space over the field $\mathbb{R}$ of real numbers. A nonnegative function $\|\cdot\|$ defined on $\mathcal{V}$ whose value at $f \in \mathcal{V}$ is denoted by $\|f\|$ is called a* **norm** *on $\mathcal{V}$ if it satisfies the following axioms:*

❶ $\|f\| = 0$ *if, and only if, $f = 0$ in $\mathcal{V}$;*
❷ $\|\lambda f\| = |\lambda|\,\|f\|$ *for all $\lambda \in \mathbb{R}$, and all f in $\mathcal{V}$;*
❸ $\|f+g\| \leq \|f\|+\|g\|$ *for all f and g in $\mathcal{V}$ (the triangle inequality).*

A linear space $\mathcal{V}$, equipped with a norm, is called a **normed linear space**.

https://doi.org/10.1017/CBO9780511801181.009 Published online by Cambridge University Press

Throughout this chapter $[a, b]$ will denote a nonempty, bounded and closed interval of $\mathbb{R}$, and (a, b) will signify a nonempty, bounded open interval of $\mathbb{R}$.

Example 8.1 *The set* $\mathrm{C}[a, b]$ *of real-valued functions* f*, defined and continuous on the interval* $[a, b]$*, is a normed linear space with norm*

$$\|f\|_\infty = \max_{x\in[a,b]} |f(x)| \,. \tag{8.1}$$

The norm $\|\cdot\|_\infty$ is called the ∞-**norm** or **maximum norm**; it can be thought of as an analogue of the ∞-norm for vectors introduced in Chapter 2. Thus, for the sake of notational simplicity, here we shall use the same symbol $\|\cdot\|_\infty$ as in Chapter 2, tacitly assuming in what follows that $\|f\|_\infty$ signifies the ∞-norm of a continuous function f, defined on a bounded closed interval of the real line (rather than the ∞-norm of an n-component vector). The choice of the interval $[a, b]$ over which the norm is taken will always be clear from the context and will not be explicitly highlighted in our notation. ◇

Example 8.2 *Suppose that* w *is a real-valued function, defined, continuous, positive and integrable on the interval* (a, b)*. The set* $\mathrm{C}[a, b]$ *of real-valued functions* f*, defined and continuous on* $[a, b]$*, is a normed linear space equipped with the norm*

$$\|f\|_2 = \left(\int_a^b w(x)|f(x)|^2 \,\mathrm{d}x\right)^{1/2} . \tag{8.2}$$

The norm $\|\cdot\|_2$ is called the **2-norm**. The function w is called a **weight function**. The assumptions on w allow for singular weight functions, such as $w\colon x \in (0, 1) \mapsto x^{-1/2}$ which is continuous, positive and integrable on the open interval $(0, 1)$, but is not continuous on the closed interval $[0, 1]$. The norm (8.2) can be thought of as an analogue of the 2-norm for vectors introduced in Chapter 2; thus, for the sake of simplicity, we use the same notation, $\|\cdot\|_2$, as there. As for the ∞-norm, we shall not explicitly indicate in our notation the interval over which the norm is taken. The implied choice of interval $[a, b]$ and weight function w will be clear from the context. ◇

The next lemma provides a comparison of the ∞-norm with the 2-norm, defined by (8.1) and (8.2), respectively, on $\mathrm{C}[a, b]$.

Lemma 8.1 *(i) Suppose that the real-valued weight function w is defined, continuous, positive and integrable on the interval (a, b). Then, for any function $f \in \mathrm{C}[a, b]$,*

$$\|f\|_2 \leq W\, \|f\|_\infty\,, \quad \text{where } W = \left[\int_a^b w(x)\mathrm{d}x\right]^{1/2}.$$

(ii) Given any two positive numbers ε (however small) and M (however large), there exists a function $f \in \mathrm{C}[a, b]$ such that

$$\|f\|_2 < \varepsilon\,, \quad \|f\|_\infty > M\,.$$

Proof The proof is left as an exercise (see Exercise 1). □

The definitions (2.33) and (2.34) of the vector norms $\|\cdot\|_\infty$ and $\|\cdot\|_2$ on $\mathbb{R}^n$ imply that

$$n^{-1/2}\|\boldsymbol{v}\|_\infty \leq \|\boldsymbol{v}\|_2 \leq n^{1/2}\|\boldsymbol{v}\|_\infty \quad \forall\, \boldsymbol{v} \in \mathbb{R}^n\,, \tag{8.3}$$

which means that, to all intents and purposes, these two norms are interchangeable.[1] Lemma 8.1 indicates that a similar chain of inequalities cannot possibly hold for the norms (8.1) and (8.2) on $\mathrm{C}[a, b]$, and the choice between them may therefore significantly influence the outcome of the analysis.

Stimulated by the first axiom of *norm*, we shall think of $f \in \mathrm{C}[a, b]$ as being well approximated by a polynomial p on $[a, b]$ if $\|f - p\|$ is small, where $\|\cdot\|$ is either $\|\cdot\|_\infty$ or $\|\cdot\|_2$ defined, respectively, by (8.1) or (8.2). In the light of Lemma 8.1, it should come as no surprise that the mathematical tools for the analysis of smallness of $\|f - p\|_\infty$ are quite different from those that ensure smallness of $\|f - p\|_2$. We have therefore chosen to discuss these two matters separately: the present chapter focuses on the ∞-norm (8.1), while Chapter 9 explores the use of the 2-norm (8.2).

Despite the fundamental differences between the norms (8.1) and (8.2) which we have alluded to above, there is a common underlying feature which is independent of the choice of norm: if *no limitation is imposed*

[1] The chain of inequalities (8.3) is, in fact, just a particular manifestation of the following general result from linear algebra. Suppose that $\mathcal{V}$ is a finite-dimensional linear space and let $\|\cdot\|'$ and $\|\cdot\|''$ be two norms on $\mathcal{V}$; then, there exist positive real numbers m and M such that

$$m\|v\|' \leq \|v\|'' \leq M\|v\|' \quad \forall\, v \in \mathcal{V}\,.$$

https://doi.org/10.1017/CBO9780511801181.009 Published online by Cambridge University Press

on the degree of the approximating polynomial p, then the approximation error $f-p$ can be made *arbitrarily small* in both norms. This is a central result in the theory of polynomial approximation and is formulated in the next theorem.

Theorem 8.1 (Weierstrass Approximation Theorem[1]) *Suppose that f is a real-valued function, defined and continuous on a bounded closed interval $[a,b]$ of the real line; then, given any $\varepsilon > 0$, there exists a polynomial p such that*

$$\|f - p\|_\infty \le \varepsilon.$$

Further, if w is a real-valued function, defined, continuous, positive and integrable on (a,b), then an analogous result holds in the 2-norm over the interval $[a,b]$ with weight function w.

This is an important theorem in classical analysis, and several proofs are known. It is evidently sufficient to consider only the interval $[0,1]$; a simple change of variable will then extend the proof to any bounded closed interval $[a,b]$. For a real-valued function f, defined and continuous on the interval $[0,1]$, Bernstein's proof uses the polynomial

$$p_n(x) = \sum_{k=0}^{n} p_{nk}(x) f(k/n)\,, \qquad x \in [0,1]\,,$$

where the **Bernstein polynomials** $p_{nk}(x)$ are defined by

$$p_{nk}(x) = \left(\begin{array}{c} n \\ k \end{array} \right) x^k (1-x)^{n-k}\,, \qquad x \in [0,1]\,.$$

It can then be shown that, for any $\varepsilon > 0$, there exists $n = n(\varepsilon)$ such that $\|f - p_n\|_\infty < \varepsilon$. The second part of the theorem is a direct consequence of this result, using part (i) of Lemma 8.1.

The details of the proof are given in Exercise 12. For an alternative proof, the reader is referred to Theorem 6.3 in M.J.D. Powell, *Approximation Theory and Methods*, Cambridge University Press, 1996.

[1] Karl Theodor Wilhelm Weierstrass (31 October 1815, Ostenfelde, Bavaria, Germany – 19 February 1897, Berlin, Germany) is frequently referred to as the father of modern mathematical analysis. He made fundamental contributions to the theory of series, functions of real variables, elliptic functions, converging infinite products, the calculus of variations, and the theory of bilinear and quadratic forms. Weierstrass' students included Cantor, Frobenius, Gegenbauer, Hölder, Hurwitz, Killing, Klein, Kneser, Sofia Kovalevskaya, Lie, Mertens, Minkowski, Mittag-Leffler, Schwarz and Stolz.

8.3 Best approximation in the ∞-norm

According to the Weierstrass Approximation Theorem any function f in $\mathrm{C}[a,b]$ can be approximated arbitrarily well from the set of *all* polynomials. Clearly, if instead of the set of all polynomials we restrict ourselves to the set of polynomials $\mathcal{P}_n$ of degree n or less, with n *fixed*, then it is no longer true that, for any $f \in \mathrm{C}[a,b]$ and any $\varepsilon > 0$, there exists $p_n \in \mathcal{P}_n$ such that

$$\|f - p_n\|_\infty < \varepsilon\,.$$

Consider, for example, the function $x \mapsto \sin x$ defined on the interval $[0,\pi]$ and fix $n = 0$; then $\|f - q\|_\infty \geq 1/2$ for any $q \in \mathcal{P}_0$, and therefore there is no q in $\mathcal{P}_0$ such that $\|f - q\|_\infty < 1/2$. A similar situation will arise if $\mathcal{P}_0$ is replaced by $\mathcal{P}_n$, with the polynomial degree n fixed.[1]

It is therefore relevant to enquire just how well a given function f in $\mathrm{C}[a,b]$ may be approximated by polynomials of a fixed degree $n \geq 0$. This question leads us to the following approximation problem.

(A) Given that $f \in \mathrm{C}[a,b]$ and $n \geq 0$, fixed, find $p_n \in \mathcal{P}_n$ such that

$$\|f - p_n\|_\infty = \inf_{q \in \mathcal{P}_n} \|f - q\|_\infty\,;$$

such a polynomial p_n is called a **polynomial of best approximation of degree n to the function f in the ∞-norm**.

The next theorem establishes the existence of a polynomial of best approximation, showing, in particular, that the infimum of $\|f - q\|_\infty$ over $q \in \mathcal{P}_n$ is attained. We shall consider the question of uniqueness of the polynomial of best approximation later on, in Theorem 8.5.

Theorem 8.2 *Given that $f \in \mathrm{C}[a,b]$, there exists a polynomial $p_n \in \mathcal{P}_n$ such that $\|f - p_n\|_\infty = \min_{q \in \mathcal{P}_n} \|f - q\|_\infty$.*

Proof Let us define the function $(c_0,\dots,c_n) \in \mathbb{R}^{n+1} \mapsto E(c_0,\dots,c_n)$ of $n+1$ real variables by

$$E(c_0,\dots,c_n) = \|f - q_n\|_\infty, \quad \text{where} \quad q_n(x) = c_0 + \cdots + c_n x^n.$$

[1] This is due to the fact that, for any fixed n, $\mathcal{P}_n$ is a closed subset of $\mathrm{C}[a,b]$; *i.e.*, if f does not belong to $\mathcal{P}_n$, there exists $\varepsilon > 0$ such that

$$\inf_{q \in \mathcal{P}_n} \|f - q\|_\infty > \varepsilon\,.$$

On the other hand, by the Weierstrass Theorem, the set of *all* polynomials is dense in $\mathrm{C}[a,b]$: any continuous function f can be represented as a limit of a uniformly convergent sequence of polynomials (of, in general, increasing degree) on $[a,b]$.

https://doi.org/10.1017/CBO9780511801181.009 Published online by Cambridge University Press

We shall first show that E is continuous; this will imply that E attains its bounds on any bounded closed set in $\mathbb{R}^{n+1}$. We shall then construct a nonempty bounded closed set $\mathcal{S} \subset \mathbb{R}^{n+1}$ such that the lower bound of E on $\mathcal{S}$ is the same as its lower bound over the whole of $\mathbb{R}^{n+1}$.

To show that E is continuous at each point $(c_0, \ldots, c_n) \in \mathbb{R}^{n+1}$, consider any $(\delta_0, \ldots, \delta_n) \in \mathbb{R}^{n+1}$ and define the polynomial $\eta_n \in \mathcal{P}_n$ by $\eta_n(x) = \delta_0 + \cdots + \delta_n x^n$. We see from the triangle inequality that

$$\begin{aligned} E(c_0 + \delta_0, \ldots, c_n + \delta_n) &= \|f - (q_n + \eta_n)\|_\infty \\ &\leq \|f - q_n\|_\infty + \|\eta_n\|_\infty \\ &= E(c_0, \ldots, c_n) + \|\eta_n\|_\infty \,. \end{aligned}$$

Now, for any given positive number ε, choose $\delta = \varepsilon/(1 + \cdots + K^n)$, where $K = \max\{|a|, |b|\}$. Consider any $(\delta_0, \ldots, \delta_n) \in \mathbb{R}^{n+1}$ such that $|\delta_i| \leq \delta$ for all $i = 0, \ldots, n$. Then,

$$\begin{aligned} E(c_0 + \delta_0, \ldots, c_n + \delta_n) - E(c_0, \ldots, c_n) &\leq \|\eta_n\|_\infty \\ &\leq \max_{x \in [a,b]} (|\delta_0| + |\delta_1||x| + \cdots + |\delta_n||x|^n) \\ &\leq \delta(1 + \cdots + K^n) \\ &= \varepsilon \,. \end{aligned} \tag{8.4}$$

Similarly,

$$\begin{aligned} E(c_0, \ldots, c_n) &= \|f - q_n\|_\infty = \|f - (q_n + \eta_n) + \eta_n\|_\infty \\ &\leq \|f - (q_n + \eta_n)\|_\infty + \|\eta_n\|_\infty \\ &\leq E(c_0 + \delta_0, \ldots, c_n + \delta_n) + \varepsilon \,, \end{aligned}$$

and therefore

$$E(c_0, \ldots, c_n) - E(c_0 + \delta_0, \ldots, c_n + \delta_n) \leq \varepsilon \,. \tag{8.5}$$

From (8.4) and (8.5) we deduce that

$$|E(c_0 + \delta_0, \ldots, c_n + \delta_n) - E(c_0, \ldots, c_n)| \leq \varepsilon$$

for all $(\delta_0, \ldots, \delta_n) \in \mathbb{R}^{n+1}$ such that $|\delta_i| \leq \delta$, $i = 0, \ldots, n$, where now $\delta = \varepsilon/(1 + \cdots + K^n)$ and $K = \max\{|a|, |b|\}$. Hence E is continuous at $(c_0, \ldots, c_n) \in \mathbb{R}^{n+1}$. Since $(c_0, \ldots, c_n)$ is an arbitrary point in $\mathbb{R}^{n+1}$, it follows that E is continuous on the whole of $\mathbb{R}^{n+1}$.

Let us denote by $\mathcal{S}$ the set of all points $(c_0, \ldots, c_n)$ in $\mathbb{R}^{n+1}$ such that $E(c_0, \ldots, c_n) \leq \|f\|_\infty + 1$. The set $\mathcal{S}$ is evidently bounded and closed in $\mathbb{R}^{n+1}$; further, $\mathcal{S}$ is nonempty since $E(0, \ldots, 0) = \|f\|_\infty \leq \|f\|_\infty + 1$, so that $(0, \ldots, 0) \in \mathcal{S}$. Hence the continuous function E attains its lower bound over the set $\mathcal{S}$; let us denote this lower bound by d and let $(c_0^*, \dots, c_n^*)$ denote the point in $\mathcal{S}$ where it is attained.

Since $(0, \dots, 0) \in \mathcal{S}$, it follows that

$$d = \min_{(c_0,\dots,c_n)\in\mathcal{S}} E(c_0, \dots, c_n) \le E(0, \dots, 0) = \|f\|_\infty .$$

According to the definition of $\mathcal{S}$,

$$E(c_0, \dots, c_n) > \|f\|_\infty + 1 \qquad \forall\, (c_0, \dots, c_n) \in \mathbb{R}^{n+1} \setminus \mathcal{S} .$$

Hence, if $(c_0, \dots, c_n) \notin \mathcal{S}$, then $E(c_0, \dots, c_n) > d + 1 > d$. Thus, the lower bound d of the function E over the set $\mathcal{S}$ is the same as the lower bound of E over all values of $(c_0, \dots, c_n) \in \mathbb{R}^{n+1}$. The lower bound d is attained at a point $(c_0^*, \dots, c_n^*)$ in $\mathcal{S}$; letting $p_n^*(x) = c_0^* + \cdots + c_n^* x^n$, we find that $d = \|f - p_n^*\|_\infty$ and therefore p_n^* is the required polynomial of best approximation of degree n to the function f in the ∞-norm. $\square$

Due to the nonconstructive nature of its proof, the last theorem does not actually tell us how to find a polynomial of best approximation of degree n for a given function $f \in \mathrm{C}[a, b]$. Therefore, our goal is now to devise a constructive characterisation of the property *'p_n is a polynomial of best approximation of degree n to the function f in the ∞-norm'*. Before doing so, however, let us simplify our terminology.

Writing the polynomial $q \in \mathcal{P}_n$ in the form

$$q_n(x) = c_0 + \cdots + c_n x^n ,$$

we want to choose the coefficients c_j, $j = 0, \dots, n$, so that they minimise the function E: $(c_0, \dots, c_n) \mapsto E(c_0, \dots, c_n)$ defined by

$$\begin{aligned} E(c_0, \dots, c_n) &= \|f - q\|_\infty \\ &= \max_{x\in[a,b]} |f(x) - c_0 - \cdots - c_n x^n| \end{aligned}$$

over $\mathbb{R}^{n+1}$. Since the polynomial of best approximation is to *minimise* (over $q \in \mathcal{P}_n$) the *maximum* absolute value of the error $f(x) - q(x)$ (over $x \in [a, b]$), it is often referred to as the **minimax polynomial**; from now on, for the sake of brevity, we shall use the latter terminology.

Before we embark on the constructive characterisation of the minimax polynomial of a continuous function, let us consider a simple example which illustrates some of its key properties.

Example 8.3 *Suppose that $f \in \mathrm{C}[0, 1]$, and that f is strictly monotonic increasing on $[0, 1]$. We wish to find the minimax polynomial p_0 of degree zero for f on $[0, 1]$.*

https://doi.org/10.1017/CBO9780511801181.009 Published online by Cambridge University Press

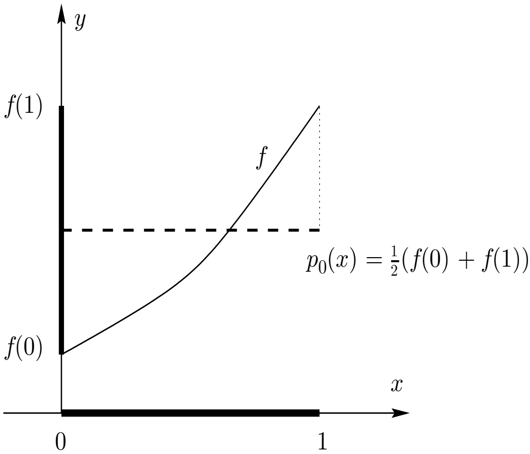


Fig. 8.1. Minimax approximation $p_0 \in \mathcal{P}_0$ of a strictly monotonic increasing continuous function f defined on the interval $[0,1]$.

The polynomial p_0 will be of the form $p_0(x) \equiv c_0$, and we need to determine $c_0 \in \mathbb{R}$ so that

$$\|f - p_0\|_\infty = \max_{x\in[0,1]} |f(x) - c_0|$$

is minimal. Since f is monotonic increasing, $f(x) - c_0$ attains its minimum at $x = 0$ and its maximum at $x = 1$; therefore $|f(x) - c_0|$ reaches its maximum value at one of the endpoints of $[0,1]$, *i.e.*,

$$E(c_0) = \max_{x\in[0,1]} |f(x) - c_0| = \max\{|f(0) - c_0|,\ |f(1) - c_0|\} .$$

Clearly,

$$E(c_0) = \begin{cases} f(1) - c_0 & \text{if } c_0 < \frac{1}{2}(f(0) + f(1)) , \\ c_0 - f(0) & \text{if } c_0 \geq \frac{1}{2}(f(0) + f(1)) . \end{cases}$$

Drawing the graph of the function $c_0 \in \mathbb{R} \mapsto E(c_0) \in \mathbb{R}$ shows that the minimum is attained when $c_0 = \frac{1}{2}(f(0) + f(1))$. Consequently, the desired minimax polynomial of degree 0 for the function f is

$$p_0(x) \equiv \tfrac{1}{2}(f(0) + f(1)) , \qquad x \in [0,1] .$$

The function f and its minimax approximation $p_0 \in \mathcal{P}_0$ are depicted in Figure 8.1.

More generally, if $f \in \mathrm{C}[a,b]$ (not necessarily monotonic), and ξ and η denote two points in $[a,b]$ where f attains its minimum and maximum values, respectively, then the minimax polynomial of degree 0 to f on $[a, b]$ is

$$p_0(x) \equiv \tfrac{1}{2}\left(f(\xi) + f(\eta)\right), \qquad x \in [a, b].$$

◇

This example shows that the minimax polynomial p_0 of degree *zero* for $f \in \mathrm{C}[a, b]$ has the property that the approximation error $f - p_0$ attains its extrema at *two* points, $x = \xi$ and $x = \eta$, with the error

$$f(x) - p_0(x) = \frac{1}{2}\left(f(x) - f(\xi)\right) + \frac{1}{2}\left(f(x) - f(\eta)\right)$$

being *negative at one point*, $x = \xi$, and *positive at the other*, $x = \eta$. We shall prove that a property of this kind holds in general; the precise formulation of the general result is given in Theorem 8.4 which is, due to the oscillating nature of the approximation error, usually referred to as the Oscillation Theorem: it gives a complete characterisation of the minimax polynomial and provides a method for its construction. We begin with a preliminary result due to de la Vallée Poussin.[1]

Theorem 8.3 (De la Vallée Poussin's Theorem) *Let $f \in \mathrm{C}[a, b]$ and $r \in \mathcal{P}_n$. Suppose that there exist $n + 2$ points $x_0 < \cdots < x_{n+1}$ in the interval $[a, b]$, such that $f(x_i) - r(x_i)$ and $f(x_{i+1}) - r(x_{i+1})$ have opposite signs, for $i = 0, \ldots, n$. Then,*

$$\min_{q \in \mathcal{P}_n} \|f - q\|_\infty \geq \min_{i=0,1,\ldots,n+1} |f(x_i) - r(x_i)|. \tag{8.6}$$

Proof The condition on the signs of $f(x_i) - r(x_i)$ is usually expressed by saying that $f - r$ has alternating signs at the points x_i, $i = 0, 1, \ldots, n+1$. Let us denote the right-hand side of (8.6) by μ. Clearly, $\mu \geq 0$; when $\mu = 0$ the statement of the theorem is trivially true, so we shall assume that $\mu > 0$. Suppose that (8.6) is false; then, for a minimax polynomial approximation $p_n \in \mathcal{P}_n$ to the function f we have[2]

$$\|f - p_n\|_\infty = \min_{q \in \mathcal{P}_n} \|f - q\|_\infty < \mu.$$

[1] Charles Jean Gustave Nicolas, Baron de la Vallée Poussin (14 August 1866, Louvain, Belgium – 2 March 1962, Louvain, Belgium) made important contributions to approximation theory and number theory, proving in 1892 that the number of primes less than n is, asymptotically as $n \to \infty$, $n/\ln n$.

[2] Recall from Theorem 8.2 that such p_n exists.

https://doi.org/10.1017/CBO9780511801181.009 Published online by Cambridge University Press

Therefore,

$$|p_n(x_i) - f(x_i)| < |r(x_i) - f(x_i)|\,, \quad i = 0, 1, \ldots, n+1\,.$$

Now,

$$r(x_i) - p_n(x_i) = [r(x_i) - f(x_i)] - [p_n(x_i) - f(x_i)]\,, \quad i = 0, 1, \ldots, n+1\,.$$

Since the first term on the right always exceeds the second term in absolute value, it follows that $r(x_i) - p_n(x_i)$ and $r(x_i) - f(x_i)$ have the same sign for $i = 0, 1, \ldots, n+1$. Hence $r - p_n$, which is a polynomial of degree n, changes sign $n+1$ times. Thus, the assumption that (8.6) is false has led to a contradiction, and the proof is complete. □

Theorem 8.3 gives a clue to formulating a constructive characterisation of the *minimax polynomial*: indeed, we shall show that if the quantities $|f(x_i) - r(x_i)|$, $i = 0, 1, \ldots, n+1$, in Theorem 8.3 are all equal to $\|f - r\|_\infty$, then $r \in \mathcal{P}_n$ is, in fact, a minimax polynomial of degree n for the function f on the interval $[a, b]$.

Theorem 8.4 (The Oscillation Theorem) *Suppose that $f \in \mathrm{C}[a, b]$. A polynomial $r \in \mathcal{P}_n$ is a minimax polynomial for f on $[a, b]$ if, and only if, there exists a sequence of $n+2$ points x_i, $i = 0, 1, \ldots, n+1$, such that $a \le x_0 < \cdots < x_{n+1} \le b$,*

$$|f(x_i) - r(x_i)| = \|f - r\|_\infty\,, \quad i = 0, 1, \ldots, n+1\,,$$

and

$$f(x_i) - r(x_i) = -[f(x_{i+1}) - r(x_{i+1})]\,, \quad i = 0, \ldots, n\,.$$

The statement of the theorem is often expressed by saying that $f - r$ attains its maximum absolute value with alternating signs at the points x_i. The points x_i, $i = 0, 1, \ldots, n+1$, in the Oscillation Theorem are referred to as **critical points**.

Proof of theorem If $f \in \mathcal{P}_n$, then the result is trivially true, with $r = f$ and any sequence of $n+2$ distinct points x_i, $i = 0, 1, \ldots, n+1$, contained in $[a, b]$. Thus, we shall suppose throughout the proof that $f \notin \mathcal{P}_n$, *i.e.*, f is such that there is no polynomial $p \in \mathcal{P}_n$ whose restriction to $[a, b]$ is identically equal to f.

The sufficiency of the condition stated in the theorem is easily shown. Suppose that the sequence of points x_i, $i = 0, 1, \ldots, n+1$, exists with the given properties. Define

$$L = \|f - r\|_\infty \qquad \text{and} \qquad E_n(f) = \min_{q \in \mathcal{P}_n} \|f - q\|_\infty .$$

From De la Vallée Poussin's Theorem, Theorem 8.3, it follows that $E_n(f) \geq L$. By the definition of $E_n(f)$ we also see that $E_n(f) \leq \|f - r\|_\infty = L$. Hence $E_n(f) = L$, and the given polynomial r is a minimax polynomial.

For the necessity of the condition, suppose that the given polynomial $r \in \mathcal{P}_n$ is a minimax polynomial for f on $[a, b]$. As $x \mapsto |f(x) - r(x)|$ is a continuous function on the bounded closed interval $[a, b]$, there exists a point in $[a, b]$ at which $|f(x) - r(x)|$ attains its maximum value, $L > 0$; let

$$x_0 = \min\{x \in [a, b]\colon |f(x) - r(x)| = L\} .$$

Now, $x_0 = b$ would imply that $|f(x) - r(x)| = L$ for all $x \in [a, b]$. As f is continuous on $[a, b]$, it would then follow that either $f(x) = r(x) + L$ for all $x \in [a, b]$ or $f(x) = r(x) - L$ for all $x \in [a, b]$; either way, we would find that $f \in \mathcal{P}_n$, which is assumed not to be the case. Therefore, $x_0 \in [a, b)$; we may assume without loss of generality that $f(x_0) - r(x_0) = L > 0$.

Now, we shall prove the existence of the next critical point, $x_1 \in (x_0, b]$ such that $f(x_1) - r(x_1) = -L$. Suppose otherwise, for contradiction; then, $-L < f(x) - r(x) \leq L$ for all x in $[a, b]$. Thus, by the continuity of f, there exists $\delta \in (0, L)$ such that $-L + \delta \leq f(x) - r(x) \leq L$ for all $x \in [a, b]$. Let us define $r^* \in \mathcal{P}_n$ by

$$r^*(x) = r(x) + \varepsilon ,$$

where $0 < \varepsilon < \min\{\delta, L\} = \delta$. Then, for all $x \in [a, b]$,

$$f(x) - r^*(x) = f(x) - r(x) - \varepsilon \geq -L + \delta - \varepsilon > -L$$

and

$$f(x) - r^*(x) = f(x) - r(x) - \varepsilon \leq L - \varepsilon < L ,$$

which means that

$$\|f - r^*\|_\infty < L = \|f - r\|_\infty .$$

Hence, $r^* \in \mathcal{P}_n$ is a better approximation to f on $[a, b]$ than $r \in \mathcal{P}_n$ is. This, however, contradicts our hypothesis that r is a polynomial of best approximation to f on $[a, b]$ from $\mathcal{P}_n$, and implies the existence of

$$x_1 = \inf\{x \in (x_0, b]\colon f(x) - r(x) = -L\} .$$

Consequently, $f(x_1) - r(x_1) = -L$ and $x_1 \in (x_0, b]$, as required; thus if $n = 0$, the proof is complete.

Let us, therefore, suppose that $n \geq 1$, and successively define the critical points

$$x_i = \inf\{x \in (x_{i-1}, b]\colon f(x) - r(x) = (-1)^i L\}, \qquad i = 1, \dots, m,$$

continuing either until $x_m = b$ or until we find an $x_m < b$ such that $|f(x) - r(x)| < L$ for all $x \in (x_m, b]$. Now, *either* $m \geq n+1$, and then the proof is complete as we will have found $n+2$ critical points, $x_0 < x_1 < \cdots < x_{n+1}$ in $[a, b]$, with the required properties, *or* $1 \leq m \leq n$.

To complete the proof of the theorem, we shall show that the second alternative, $1 \leq m \leq n$, leads to a contradiction, and is, therefore, not possible. Let us suppose, for this purpose, that $1 \leq m \leq n$, and let $\eta_0 = a$. Further, observe that, due to the definition of the points x_i, $i = 0, 1, \dots, m$,

$$\exists \eta_i \in (x_{i-1}, x_i) \quad \forall\, x \in [\eta_i, x_i) \quad |f(x) - r(x)| < L, \quad i = 1, \dots, m,$$

and define $\eta_{m+1} = b$.

It follows from the choice of the η_i, $i = 0, 1, \dots, m+1$, that the following properties hold:

(a) $|f(x) - r(x)| \leq L$ for all $x \in [\eta_i, \eta_{i+1}]$ and all $i = 0, 1, \dots, m$;
(b) for each $i = 0, 1, \dots, m$ there exists $x \in [\eta_i, \eta_{i+1}]$ (say, $x = x_i$), such that $f(x) - r(x) = (-1)^i L$;
(c) there exist no $i \in \{0, 1, \dots, m\}$ and $x \in [\eta_i, \eta_{i+1}]$ such that $f(x) - r(x) = (-1)^{i+1} L$;
(d) $|f(\eta_i) - r(\eta_i)| < L$ for all $i = 1, \dots, m$.

Now, let

$$v(x) = \prod_{i=1}^{m} (\eta_i - x),$$

and define

$$r^*(x) = r(x) + \varepsilon v(x),$$

where $\varepsilon > 0$ is a fixed real number, to be chosen below. Since, by hypothesis, $1 \leq m \leq n$, it follows that $r^* \in \mathcal{P}_n$. Let us consider the behaviour of the difference

$$f(x) - r^*(x) = f(x) - r(x) - \varepsilon v(x)$$

on each of the intervals $[\eta_i, \eta_{i+1}]$, $i = 0, 1, \ldots, m$ (whose union is $[a, b]$). We shall prove that, for $\varepsilon > 0$ sufficiently small,

$$|f(x) - r^*(x)| < L = \|f - r\|_\infty$$

for all x in $[\eta_i, \eta_{i+1}]$ and all $i = 0, 1, \ldots, m$; *i.e.*, $\|f - r^*\|_\infty < \|f - r\|_\infty$, contradicting the fact that $r \in \mathcal{P}_n$ is a minimax polynomial for f on $[a, b]$, and refuting the hypothesis that $1 \le m \le n$.

Take, for example, the interval $[\eta_0, \eta_1]$. For each x in $[\eta_0, \eta_1)$ we have $v(x) > 0$ and therefore, by the definition of $r^*(x)$ and property (a) above,

$$f(x) - r^*(x) \le L - \varepsilon v(x) < L\,, \qquad x \in [\eta_0, \eta_1).$$

Further, as $v(\eta_1) = 0$, it follows from (d) that

$$f(\eta_1) - r^*(\eta_1) = f(\eta_1) - r(\eta_1) < L\,.$$

Therefore, $f(x) - r^*(x) < L$ for each x in $[\eta_0, \eta_1]$. For a lower bound on $f(x) - r^*(x)$, note that by (a) and (c), $f(x) - r(x) > -L$ for all x in $[\eta_0, \eta_1]$. As $f - r$ is a continuous function on $[\eta_0, \eta_1]$, there exists $\delta_1 \in (0, L)$ such that $f(x) - r(x) \ge -L + \delta_1$ for all x in $[\eta_0, \eta_1]$. Thus, for $0 < \varepsilon < \min\{L, \delta_1, \varepsilon_1\}$, where

$$\varepsilon_1 = \frac{\delta_1}{\max_{x \in [\eta_0, \eta_1]} |v(x)|}\,,$$

we have that

$$f(x) - r^*(x) \ge -L + \delta_1 - \varepsilon |v(x)| > -L\,, \qquad x \in [\eta_0, \eta_1)\,.$$

Further, by (d) above,

$$f(\eta_1) - r^*(\eta_1) = f(\eta_1) - r(\eta_1) > -L\,.$$

Hence, $f(x) - r^*(x) > -L$ for all $x \in [\eta_0, \eta_1]$, for $0 < \varepsilon < \min\{L, \delta_1, \varepsilon_1\}$. Combining the upper and lower bounds on $f(x) - r^*(x)$, we deduce that

$$|f(x) - r^*(x)| < L = \|f - r\|_\infty\,, \qquad x \in [\eta_0, \eta_1]\,.$$

Arguing in the same manner on each of the other intervals $[\eta_i, \eta_{i+1}]$, $i = 1, \ldots, m$, with $0 < \varepsilon < \min\{L, \delta_{i+1}, \varepsilon_{i+1}\}$, $i = 1, \ldots, m$, and δ_{i+1} and ε_{i+1} defined analogously to δ_1 and ε_1 above, we conclude that

$$|f(x) - r^*(x)| < L = \|f - r\|_\infty\,, \qquad x \in [\eta_i, \eta_{i+1}]\,, \quad i = 0, 1, \ldots, m\,,$$

and hence, for $0 < \varepsilon < \min\{L, \delta_1, \varepsilon_1, \ldots, \delta_{m+1}, \varepsilon_{m+1}\}$,

$$\|f - r^*\|_\infty < L = \|f - r\|_\infty\,.$$

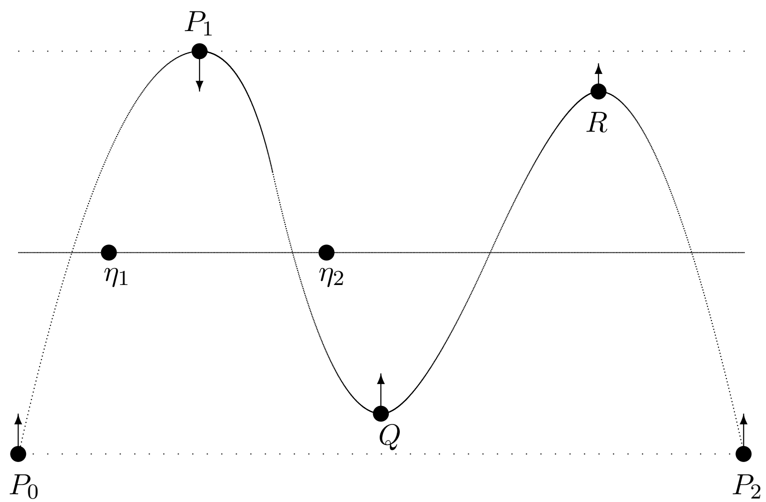


Fig. 8.2. The Oscillation Theorem: the difference $f(x) - r(x)$, where r is a cubic approximation to a continuous function f, and the effect of replacing $r(x)$ by $r^*(x) = r(x) - \varepsilon v(x)$, where $v(x) = (\eta_1 - x)(\eta_2 - x)$.

As r^* is in $\mathcal{P}_n$, the last inequality contradicts our assumption that r is a polynomial of best approximation to f on $[a, b]$ from $\mathcal{P}_n$. The contradiction rules out the possibility that $1 \leq m \leq n$. Since $m \geq 1$, it follows that $m \geq n + 1$, and the proof is complete. □

In the proof of the Oscillation Theorem we supposed, without loss of generality, that $f(x_0) - r(x_0) = L > 0$, where $L = \|f - r\|_\infty$. When $f(x_0) - r(x_0) = -L < 0$ the proof is analogous, except we then define $r^*(x) = r(x) - \varepsilon$ to prove the existence of the critical point $x_1 \in (x_0, b]$ and, in the discussion of the case $1 \leq m \leq n$, we let

$$r^*(x) = r(x) - \varepsilon v(x) \,,$$

with $v(x)$ and $\varepsilon > 0$ defined as before.

A typical situation is illustrated in Figure 8.2, which represents the difference $f - r$, where r is a polynomial approximation of degree 3 to a continuous function f. Here $|f - r|$ attains its maximum value with alternate signs at the points P_0, P_1 and P_2, so that $m = 2 < n = 3$. Let x_0, x_1 and x_2 denote the x-coordinates of P_0, P_1, P_2, respectively. Clearly, $f(x_0) - r(x_0) = -L < 0$, where $L = \|f - r\|_\infty$. Also, the two points η_1 and η_2 are as shown, $v(x) = (\eta_1 - x)(\eta_2 - x)$, and the effect of replacing r by $r^*(x) = r(x) - \varepsilon v(x)$, with $\varepsilon > 0$, is indicated by the arrows. Since $f - r^* = f - r + \varepsilon v(x)$ and v is negative for $x \in (\eta_1, \eta_2)$ and positive outside (η_1, η_2), $|f - r^*|$ will be smaller than $|f - r|$ at each of the points P_i, $i = 0, 1, 2$. There are two other local extrema for the error function $f - r$: a minimum at Q and a maximum at R. Since both these points are to the right of η_2, where $v(x) > 0$, we shall have $f - r^* > f - r$ at both of Q and R, and $|f - r^*| > |f - r|$ at R. The magnitude of the extra term $\varepsilon v(x)$ must therefore be limited by the need to avoid the new difference $f - r^*$ becoming too large at R. We can achieve this by selecting $\varepsilon > 0$ sufficiently small. In this illustration the polynomial $r \in \mathcal{P}_3$ is not a minimax approximation to f on the given interval, since we can construct a better approximation r^* which is also in $\mathcal{P}_3$.

We can now apply the Oscillation Theorem to prove that the minimax polynomial is unique.

Theorem 8.5 (Uniqueness Theorem) *Suppose that* $[a, b]$ *is a bounded closed interval of the real line. Each* $f \in \mathrm{C}[a, b]$ *has a unique minimax polynomial* $p_n \in \mathcal{P}_n$ *on* $[a, b]$.

Proof Suppose that $q_n \in \mathcal{P}_n$ is also a minimax polynomial for f, and that p_n and q_n are distinct. Then,

$$\|f - p_n\|_\infty = \|f - q_n\|_\infty = E_n(f)\,,$$

where, as in the proof of the Oscillation Theorem, we have used the notation

$$E_n(f) = \min_{q \in \mathcal{P}_n} \|f - q\|_\infty\,.$$

This implies, by the triangle inequality, that

$$\begin{aligned}\|f - \tfrac{1}{2}(p_n + q_n)\|_\infty &= \|\tfrac{1}{2}(f - p_n) + \tfrac{1}{2}(f - q_n)\|_\infty \\ &\le \tfrac{1}{2}\|f - p_n\|_\infty + \tfrac{1}{2}\|f - q_n\|_\infty \\ &= E_n(f)\,.\end{aligned}$$

Therefore $\frac{1}{2}(p_n + q_n) \in \mathcal{P}_n$ is also a minimax polynomial approximation to f on $[a, b]$. By the Oscillation Theorem there exists a sequence of $n + 2$ critical points x_i, $i = 0, 1, \ldots, n + 1$, at which

$$\left|f(x_i) - \tfrac{1}{2}\left(p_n(x_i) + q_n(x_i)\right)\right| = E_n(f), \quad i = 0, 1, \ldots, n + 1\,.$$

This is equivalent to

$$| \, (f(x_i) - p_n(x_i)) \, + \, (f(x_i) - q_n(x_i)) \, | = 2E_n(f) \, .$$

Now

$$|f(x_i) - p_n(x_i)| \le \max_{x \in [a,b]} |f(x) - p_n(x)| = \|f - p_n\|_\infty = E_n(f) \, ,$$

and, for the same reason,

$$|f(x_i) - q_n(x_i)| \le E_n(f) \, .$$

It therefore follows[1] that

$$f(x_i) - p_n(x_i) = f(x_i) - q_n(x_i) \, , \quad i = 0, 1, \ldots, n+1 \, .$$

Thus, the difference $p_n - q_n$ vanishes at $n+2$ distinct points. As $p_n - q_n$ is a polynomial of degree n or less, it follows that $p_n - q_n$ is identically zero. This, however, contradicts our initial hypothesis that p_n and q_n are distinct, and implies the uniqueness of the minimax polynomial $p_n \in \mathcal{P}_n$ for $f \in \mathrm{C}[a,b]$. □

As an application of the Oscillation Theorem, we consider the construction of the minimax approximation $p_1 \in \mathcal{P}_1$ of degree 1 to a function $f \in \mathrm{C}[a,b]$ on the interval $[a,b]$, where we assume that f has a continuous and strictly monotonic increasing derivative f' on this interval.

We seek the minimax polynomial $p_1 \in \mathcal{P}_1$ in the form $p_1(x) = c_1 x + c_0$. The difference $f(x) - (c_1 x + c_0)$ attains its extrema either at the endpoints of the interval $[a,b]$ or at points where its derivative $f'(x) - c_1$ is zero. Since f' is strictly monotonic increasing it can only take the value c_1 at one point at most. Therefore the endpoints of the interval, a and b, are critical points. Let us denote by d the third critical point whose location inside (a,b) remains to be determined. Since the critical point $x = d$ is an internal extremum of $f(x) - (c_1 x + c_0)$, it follows that

$$(f(x) - (c_1 x + c_0))'|_{x=d} = 0 \, .$$

By the Oscillation Theorem, with $x_0 = a$, $x_1 = d$, $x_2 = b$, we have the

[1] We use the following elementary result: if P and Q are two real numbers and E is a nonnegative real number such that $|P + Q| = 2E$, $|P| \le E$ and $|Q| \le E$, then $P = Q$. This follows by noting that $(P - Q)^2 = 2P^2 + 2Q^2 - (P+Q)^2 \le 2E^2 + 2E^2 - 4E^2 = 0$, and hence $P - Q = 0$. In the proof of the theorem we apply this with $P = f(x_i) - p_n(x_i)$, $Q = f(x_i) - q_n(x_i)$ and $E = E_n(f)$.

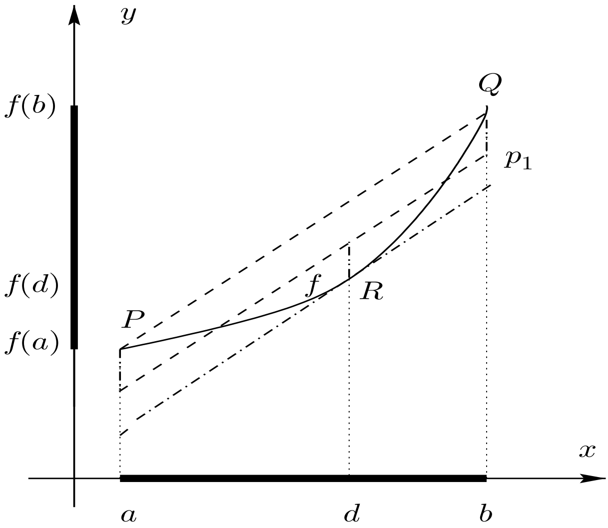


Fig. 8.3. Construction of minimax polynomial of degree 1.

equations

$$\left.\begin{aligned} f(a) - (c_1 a + c_0) &= A\,, \\ f(d) - (c_1 d + c_0) &= -A\,, \\ f(b) - (c_1 b + c_0) &= A\,, \end{aligned}\right\} \tag{8.7}$$

where either $A = L$ or $A = -L$, with $L = \max_{x\in[a,b]} |f(x) - p_1(x)|$. Along with the condition

$$f'(d) = c_1 \tag{8.8}$$

this gives four equations to determine the unknowns d, c_1, c_0 and A.

Subtracting the first equation in (8.7) from the third equation, we get $f(b) - f(a) = c_1(b - a)$, whereby $c_1 = (f(b) - f(a))/(b - a)$. Now, by the Mean Value Theorem, Theorem A.3, with this choice of c_1 equation (8.8) has at least one solution, d, in the open interval in (a, b). In fact, the value of d is uniquely determined by (8.8), as f' is continuous and strictly monotonic increasing. Next, c_0 can be determined by adding the second equation in (8.7) to the first. Having calculated both c_1 and c_0 we insert them into the first equation in (8.7) to obtain A; finally $L = |A|$.

The construction of the minimax polynomial p_1 is illustrated in Figure 8.3; R is the point at which the tangent to the curve $y = f(x)$ is parallel to the chord PQ; the graph of $p_1(x)$ is parallel to these two lines, and lies half-way between them.

Table 8.1. *The first seven Chebyshev Polynomials:* $T_0, T_1, \ldots, T_6$.

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

8.4 Chebyshev polynomials

There are very few functions for which it is possible to write down in simple closed form the minimax polynomial. One such problem of practical importance concerns the approximation of a power of x by a polynomial of lower degree. The minimax approximation in this case is given in terms of Chebyshev polynomials.[1]

Definition 8.2 *The* **Chebyshev polynomial** T_n *of degree* n *is defined, for* $x \in [-1, 1]$, *by*

$$T_n(x) = \cos(n \cos^{-1} x)\,, \qquad n = 0, 1, 2, \ldots\,.$$

Despite its unusual form, T_n is a polynomial in disguise. For example, $T_0(x) \equiv 1$, $T_1(x) = x$ for all $x \in [-1, 1]$, and so on. In order to show that this is true in general, we recall the trigonometric identity

$$\cos{(n+1)\vartheta} + \cos{(n-1)\vartheta} = 2\cos\vartheta\cos n\vartheta\,,$$

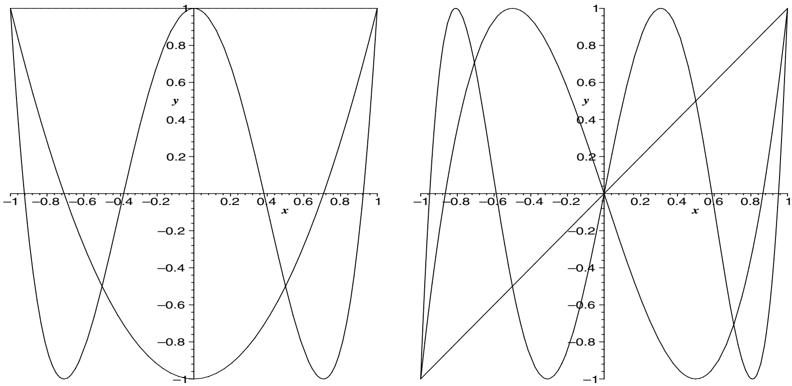
and set $\vartheta = \cos^{-1} x$, with $x \in [-1, 1]$, to obtain the recurrence relation

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x)\,, \qquad n = 1, 2, 3, \ldots\,, \quad x \in [-1, 1]\,.$$

Since T_0 and T_1 have already been shown to be polynomials on $[-1, 1]$, we deduce from this recurrence relation, by induction, that T_n is a polynomial of degree n on $[-1, 1]$ for each $n \geq 0$. A list of the first seven Chebyshev polynomials is given in Table 8.1.

[1] Pafnuty Lvovich Chebyshev (16 May 1821, Okatovo, Russia – 8 December 1894, St Petersburg, Russia). In 1850 Chebyshev proved the Bertrand conjecture, that there is always at least one prime between n and $2n$ for $n \geq 2$. He also came close to proving the Prime Number Theorem which states that the number of primes less than n is, asymptotically as $n \to \infty$, $n/\ln n$. The proof was completed, independently, by Dirichlet and de la Vallée Poussin two years after Chebyshev's death. Chebyshev made important contributions to probability theory, orthogonal functions and the theory of integrals.

https://doi.org/10.1017/CBO9780511801181.009 Published online by Cambridge University Press



(a) (b)

Fig. 8.4. The first three Chebyshev polynomials of (a) even degree, T_0, T_2, T_4, and (b) odd degree T_1, T_3, T_5, plotted on the interval $[-1, 1]$.

The polynomials T_0, T_2, T_4, and T_1, T_3, T_5, are depicted in Figure 8.4. We see that the even-degree Chebyshev polynomials are even functions; (*i.e.*, $T_{2k}(-x) = T_{2k}(x)$ for all $x \in [-1, 1]$) and the odd-degree ones are odd functions (*i.e.*, $T_{2k+1}(-x) = -T_{2k+1}(x)$ for all $x \in [-1, 1]$). They all map the interval $[-1, 1]$ into itself.[1]

The proof of the next lemma is straightforward and is left as an exercise (see Exercise 10).

Lemma 8.2 *The Chebyshev polynomials have the following properties:*

(i) $T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x)$, $x \in [-1, 1]$, $n = 1, 2, 3, \ldots$;

(ii) *for* $n \geq 1$, T_n *is a polynomial in* x *of degree* n *on the interval* $[-1, 1]$, *with leading coefficient* $2^{n-1}x^n$;

(iii) T_n *is an even function on* $[-1, 1]$ *if* n *is even, and an odd function on* $[-1, 1]$ *if* n *is odd,* $n \geq 0$;

(iv) *for* $n \geq 1$, *the zeros of* T_n *are at*

$$x_j = \cos\frac{(2j-1)\pi}{2n}, \qquad j = 1, \ldots, n;$$

[1] In Maple, typing `plot(orthopoly[T](7,x), x=-1..1, y=-1..1);` will, for example, plot the graph of the Chebyshev polynomial T_7 of degree 7 in x; T_8, T_9, etc., can be obtained similarly. Incidentally, you may be wondering why T_n and not C_n is used to denote the Chebyshev polynomial of degree n. The reasons are largely historical: in some older books and articles Chebyshev's Russian surname has been transliterated from the Cyrillic original as Tchebyshev, following the French and German transliterations Tchebychef and Tschebyscheff, respectively.

they are all real and distinct, and lie in $(-1,1)$*;*

(v) $|T_n(x)| \le 1$ *for all* $x \in [-1,1]$ *and all* $n \ge 0$*;*

(vi) *for* $n \ge 1$, $T_n(x) = \pm 1$, *alternately at the* $n+1$ *points* $x_k = \cos(k\pi/n)$, $k = 0, 1, \dots, n$.

We can now apply the Oscillation Theorem to construct the minimax polynomial of degree n for $f\colon x \mapsto x^{n+1}$ on the interval $[-1,1]$.

Theorem 8.6 *Suppose that* $n \ge 0$. *The polynomial* $p_n \in \mathcal{P}_n$ *defined by*

$$p_n(x) = x^{n+1} - 2^{-n} T_{n+1}(x)\,, \qquad x \in [-1,1]\,,$$

is the minimax approximation of degree n *to the function* $x \mapsto x^{n+1}$ *on the interval* $[-1,1]$.

Proof By part (ii) of Lemma 8.2, $p_n \in \mathcal{P}_n$. Since

$$x^{n+1} - p_n(x) = 2^{-n} T_{n+1}(x)\,,$$

by parts (v) and (vi) of Lemma 8.2, the difference $x^{n+1} - p_n(x)$ does not exceed 2^{-n} in the interval $[-1,1]$, and attains this value with alternating signs at the $n+2$ points $x_k = \cos(k\pi/(n+1))$, $k = 0, 1, \dots, n+1$. Therefore, by the Oscillation Theorem, p_n is the (unique) minimax polynomial approximation from $\mathcal{P}_n$ to the function $x \mapsto x^{n+1}$ over $[-1,1]$. $\square$

A polynomial of degree n whose leading coefficient, the coefficient of x^n, is equal to 1, is called a **monic polynomial** of degree n. For example, the polynomial $r \in \mathcal{P}_{n+1}$ defined by $r(x) = x^{n+1} - q(x)$ with $q \in \mathcal{P}_n$, is a monic polynomial of degree $n+1$.

Corollary 8.1 *Suppose that* $n \ge 0$. *Among all monic polynomials of degree* $n+1$ *the polynomials* $2^{-n}T_{n+1}$ *and* $-2^{-n}T_{n+1}$ *have the smallest* ∞*-norm on the interval* $[-1,1]$.

Proof Let $\mathcal{P}^1_{n+1}$ denote the set of all monic polynomials of degree $n+1$. Any $r \in \mathcal{P}^1_{n+1}$ can be regarded as the difference between the function $x \mapsto x^{n+1}$ and a polynomial of lower degree, *i.e.*, $r(x) = x^{n+1} - q(x)$ with $q \in \mathcal{P}_n$. Hence, by Theorem 8.6,

$$\begin{aligned}
\min_{r \in \mathcal{P}^1_{n+1}} \|r\|_\infty &= \min_{q \in \mathcal{P}_n} \|x^{n+1} - q\|_\infty \\
&= \|x^{n+1} - (x^{n+1} - 2^{-n}T_{n+1})\|_\infty \\
&= \|2^{-n}T_{n+1}\|_\infty\,;
\end{aligned}$$

https://doi.org/10.1017/CBO9780511801181.009 Published online by Cambridge University Press

the minimum is, therefore, achieved when $r \in \mathcal{P}^1_{n+1}$ is one of the monic polynomials $2^{-n}T_{n+1}$ or $-2^{-n}T_{n+1}$.

□

8.5 Interpolation

We close the body of this chapter with another application of Chebyshev polynomials: it concerns the 'optimal' choice of interpolation points in Lagrange interpolation. In Chapter 6 the error between an $n+1$ times continuously differentiable function f, defined on a closed interval $[a, b]$ of the real line, and its Lagrange interpolation polynomial p_n of degree n, $n \geq 0$, with interpolation points $\xi_0, \ldots, \xi_n$, was shown to have the form

$$f(x) - p_n(x) = \frac{f^{(n+1)}(\eta)}{(n+1)!}\,\pi_{n+1}(x)\,, \tag{8.9}$$

where $\eta = \eta(x) \in (a, b)$ and

$$\pi_{n+1}(x) = (x - \xi_0)\ldots(x - \xi_n)\,. \tag{8.10}$$

Clearly, π_{n+1} is a monic polynomial of degree $n+1$.

In a practical application the values ξ_i and $f(\xi_i)$, $i = 0, 1, \ldots, n$, may be already given. However, in a situation where $[a, b] = [-1, 1]$ and the ξ_i, $i = 0, 1, \ldots, n$, can be freely chosen in the interval $[-1, 1]$, Corollary 8.1 suggests that they should be taken as the zeros of the Chebyshev polynomial T_{n+1}, for then π_{n+1} will have the smallest ∞-norm on the interval $[-1, 1]$ among all monic polynomials. This observation motivates the following result.

Theorem 8.7 *Suppose that f is a real-valued function, defined and continuous on the closed real interval $[a, b]$, and such that the derivative of f of order $n+1$ is continuous on $[a, b]$. Let $p_n \in \mathcal{P}_n$ denote the Lagrange interpolation polynomial of f, with interpolation points*

$$\xi_j = \tfrac{1}{2}(b-a)\cos\frac{(j+\frac{1}{2})\pi}{n+1} + \tfrac{1}{2}(b+a)\,, \quad j = 0, 1, \ldots, n\,;$$

then

$$\|f - p_n\|_\infty \leq \frac{(b-a)^{n+1}}{2^{2n+1}(n+1)!}M_{n+1}$$

where $M_{n+1} = \max_{\zeta \in [a,b]} |f^{(n+1)}(\zeta)|$.

Proof Let $\tau_j = \cos\big((j+\frac{1}{2})\pi/(n+1)\big)$, $j = 0, 1, \dots, n$, denote the zeros of the polynomial $T_{n+1}(t)$ (in the interval $(-1, 1)$). Hence,

$$\prod_{j=0}^{n}(t-\tau_j) = 2^{-n}T_{n+1}(t)\,, \qquad t \in [-1, 1]\,.$$

Let us define the points ξ_j, $j = 0, 1, \dots, n$, as in the statement of the theorem. Clearly $\xi_j \in (a, b)$ is the image of $\tau_j \in (-1, 1)$ under the linear transformation $t \mapsto x = \frac{1}{2}(b-a)t + \frac{1}{2}(b+a)$; we note in passing that the inverse of this mapping is $x \mapsto t(x) = (2x-a-b)/(b-a)$; thus,

$$\prod_{j=0}^{n}(x-\xi_j) = \left(\frac{b-a}{2}\right)^{n+1}\prod_{j=0}^{n}(t(x)-\tau_j) = \left(\frac{b-a}{2}\right)^{n+1} 2^{-n}T_{n+1}(t(x))\,.$$

The required bound now follows from (8.9), since $|T_{n+1}(t(x))| \le 1$ for all $x \in [a, b]$, and therefore $|\pi_{n+1}(x)| \le (b-a)^{n+1}2^{-2n-1}$. □

The De la Vallée Poussin Theorem, Theorem 8.3, suggests the notion of a **near-minimax** polynomial, which is a polynomial $p_n \in \mathcal{P}_n$ such that the difference $f(x) - p_n(x)$ changes sign at $n+1$ points ξ_j, $j = 0, 1, \dots, n$, with $a < \xi_0 < \cdots < \xi_n < b$; for the difference $f(x) - p_n(x)$ then attains a local maximum or minimum with alternating signs in each of the intervals $[a, \xi_0), (\xi_0, \xi_1), \dots, (\xi_n, b]$. The positions of these alternating local maxima and minima are then the points x_i, $i = 0, 1, \dots, n+1$, required by Theorem 8.3, and we therefore know that the ∞-norm of the error of the minimax polynomial lies between the least and greatest of the absolute values of these local maxima and minima. In particular, we should expect that if the sizes of these local maxima and minima are not greatly different, then the error of the near-minimax approximation should not be very much larger than the error of the minimax approximation.

Given any set of points ξ_i, $i = 0, 1, \dots, n$, with $a < \xi_0 < \cdots < \xi_n < b$, the polynomial $\pi_{n+1}(x) = (x-\xi_0)\dots(x-\xi_n)$ changes sign at the $n+1$ points ξ_j, $j = 0, 1, \dots, n$. Let us assume that $f \in \mathrm{C}[a, b]$, $f^{(n+1)}$ exists and is continuous on $[a, b]$, and $f^{(n+1)}$ has the same sign on the whole of (a, b). It then follows that the product $f^{(n+1)}(\eta)\pi_{n+1}(x)$ has exactly $n+1$ sign-changes in the open interval (a, b) for any $\eta \in (a, b)$. Thus, according to (8.9), the Lagrange interpolation polynomial p_n of degree n for the function f, with interpolation points ξ_j, $j = 0, 1, \dots, n$, contained in the open interval (a, b), is a near-minimax polynomial from $\mathcal{P}_n$ for f on $[a, b]$.

We have therefore just shown that if $f^{(n+1)}$ exists and is continuous on the closed interval $[a, b]$, and has the same sign on the open interval

https://doi.org/10.1017/CBO9780511801181.009 Published online by Cambridge University Press

(a, b), then the polynomial constructed by interpolating at the points ξ_j, $j = 0, 1, \ldots, n$, obtained by linearly mapping the $n+1$ zeros of the Chebyshev polynomial $T_{n+1}(t)$ from $(-1, 1)$ to (a, b), is a near-minimax approximation from $\mathcal{P}_n$ for the function $f \in \mathrm{C}[a, b]$ on the interval $[a, b]$. Notice that if we use equally spaced interpolation points, so that $\xi_j = a + j(b-a)/n$, $j = 0, 1, \ldots, n$, $n \geq 1$, we shall not obtain a near-minimax approximation, since the interpolation error now changes sign at only $n-1$ points, the interpolation points which are internal to (a, b).

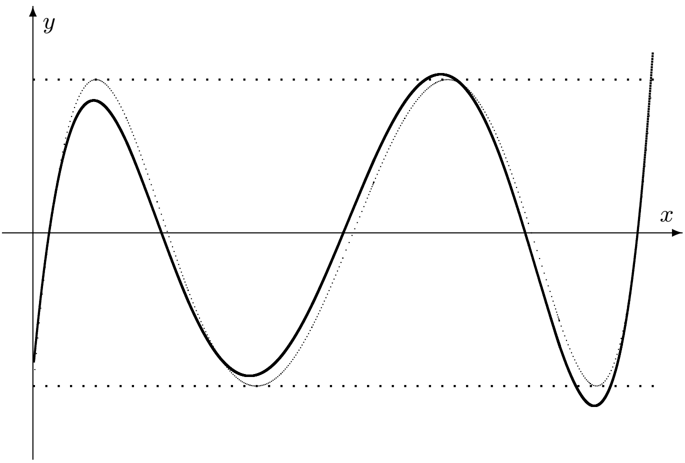


Fig. 8.5. Comparison of two polynomial approximations to e^{2x} on [0,1]: the thick curve is the error of the minimax approximation; the thin curve is the error of the polynomial obtained by interpolation at the Chebyshev points.

As an illustration, Figure 8.5 shows the errors of two approximations of degree 4 to the function $f(x) = \mathrm{e}^{2x}$ over the interval $[0, 1]$. One of these is the minimax approximation, and the other is obtained by interpolation at the zeros of $T_5(t)$. It is clear that they are quite close; in fact the ∞-norms of the errors are 0.0015 and 0.0017 respectively.

In the next chapter we shall show that the least squares polynomial approximation to a continuous real-valued function is also near-minimax in this sense.

An alternative and very easy way of constructing polynomial approximations to many simple, smooth, functions is to truncate their Taylor series expansion. For example,

$$\mathrm{e}^{kx} = 1 + kx + \cdots + \frac{k^n x^n}{n!} + \cdots,$$

https://doi.org/10.1017/CBO9780511801181.009 Published online by Cambridge University Press

so we obtain a polynomial approximation $p_n(x)$ by taking the terms of this series up to the one involving x^n. Then, clearly,

$$\mathrm{e}^{kx} - p_n(x) = \sum_{r=n+1}^{\infty} \frac{k^r x^r}{r!}\,.$$

Over the interval $[0,1]$, for example, this difference is nonnegative and monotonic increasing; it does not change sign at all. Hence the polynomial $p_n \in \mathcal{P}_n$ thus constructed is quite certainly not a near-minimax approximation for $x \mapsto \mathrm{e}^{kx}$ on $[0,1]$. Nevertheless, $\max_{x\in[0,1]} |\mathrm{e}^{kx} - p_n(x)|$ can be made arbitrarily small by choosing n sufficiently large.

8.6 Notes

For further details on the topics presented in this chapter, we refer to

- M.J.D. Powell, *Approximation Theory and Methods*, Cambridge University Press, Cambridge, 1996.

The Weierstrass Theorem is discussed in Chapter 6 of that book, and is stated in its Theorem 6.3. Although the proof presented by Powell uses the Bernstein polynomials, it is different from the more elementary but slightly lengthier argument proposed in Exercise 12 here: it relies on a proof of Bohman and Korovkin based on properties of monotone operators; see, also, p. 66 in Chapter 3 of

- E.W. Cheney, *Introduction to Approximation Theory*, McGraw–Hill, New York, 1966.

The notes contained on pp. 224–233 of Cheney's book are particularly illuminating.

The proof of the Weierstrass Theorem as proposed in Exercise 12, including the definition of what we today call Bernstein polynomials, stem from a paper of Sergei Natanovich Bernstein (1880–1968), entitled 'Démonstration du théorème de Weierstrass fondée sur le calcul des probabilités', *Comm. Soc. Math. Kharkow* **13**, 1–2, 1912/13.

Weierstrass' main contributions to approximation theory, as well as those of other mathematicians (including Picard, Volterra, Runge, Lebesgue, Mittag-Leffler, Fejér, Landau, de la Vallée Poussin, Bernstein), are reviewed in the extensive historical survey by Allan Pinkus, Weierstrass and approximation theory, *J. Approx. Theory* **107**, 1–66, 2000. Further details about the history of the subject can be found at the history of approximation theory website maintained by Allan Pinkus and Carl de Boor: `http://www.cs.wisc.edu/~deboor/HAT/`

The second part of Theorem 8.1 concerning the approximability of a continuous function by polynomials in the 2-norm is not usually presented as part of the classical Weierstrass Theorem which is posed in the ∞-norm. Here, we have chosen to state these results together in order to highlight the analogy, as well as to motivate the use of the 2-norm in polynomial approximation in the next chapter, Chapter 9.

In both Cheney's and Powell's books minimax approximation is treated in the more general framework of Haar systems. An $(n+1)$-dimensional linear subspace $\mathcal{A}$ of $\mathrm{C}[a, b]$ is said to satisfy the *Haar condition* if, for every nonzero p in $\mathcal{A}$, the number of roots of the equation $p(x) = 0$ in the interval $[a, b]$ is less than $n + 1$. The concept of Haar system is due to Alfred Haar (1885–1933), Die Minkowskische Geometrie und die Annäherung an stetige Funktionen, *Math. Ann.* **78**, 294–311, 1918; this paper contains Haar's Theorem which characterises finite-dimensional Haar systems in spaces of continuous functions. The *Characterisation Theorem*, formulated as Theorem 7.2 in Powell's book, shows that the Oscillation Theorem, Theorem 8.4 of the present chapter, remains valid in a more general setting when the set of polynomials $\{1, x, \ldots, x^n\}$ is replaced by an $(n + 1)$-dimensional Haar system of functions contained in $\mathrm{C}[a, b]$.

Exercises

8.1 Give a proof of Lemma 8.1.

8.2 Suppose that the real-valued function f is continuous and even on the interval $[-a, a]$, that is, $f(x) = f(-x)$ for all $x \in [-a, a]$. By using the Uniqueness Theorem, or otherwise, show that the minimax polynomial approximation of degree n is an even function. Deduce that the minimax polynomial approximation of degree $2n$ is also the minimax polynomial approximation of degree $2n+1$. What does this imply about the sequence of critical points for the minimax polynomial p_{2n}?

8.3 State and prove similar results to those in Exercise 2, for the case where f is an odd function, that is, $f(x) = -f(-x)$ for all $x \in [-a, a]$.

8.4 (i) Construct the minimax polynomial $p_2 \in \mathcal{P}_2$ on the interval $[-1, 1]$ for the function g defined by $g(x) = \sin x$.

(ii) Construct the minimax polynomial $p_3 \in \mathcal{P}_3$ on the interval $[-1, 1]$ for the function h defined by $h(x) = \cos x^2$.
(Use the results of Exercises 2 and 3.)

8.5 The function H is defined by $H(x) = 1$ if $x > 0$, $H(x) = -1$ if $x < 0$, and $H(0) = 0$. Show that for any $n \geq 0$ and any $p_n \in \mathcal{P}_n$, $\|H - p_n\|_\infty \geq 1$ on the interval $[-1, 1]$. Construct the polynomial, of degree 0, of best approximation to H on the interval $[-1, 1]$, and show that it is unique. (Note that since H is discontinuous most of the theorems in this chapter are not applicable.)

Show that the polynomial of best approximation, of degree 1, to H on $[-1, 1]$ is not unique, and give an expression for its most general form.

8.6 Suppose that $t_1 < t_2 < \cdots < t_k$ are k distinct points in the interval $[a, b]$; for any function f defined on $[a, b]$, write $Z_k(f) = \max_{i=1}^k |f(x_i)|$. Explain why $Z_k(\,\cdot\,)$ is not a norm on the space of functions which are continuous on $[a, b]$; show that it is a norm on the space of polynomials of degree n, provided that $k > n$.

In the case $k = 3$, with $t_1 = 0$, $t_2 = \frac{1}{2}$, $t_3 = 1$, where we wish to approximate the function $f\colon x \mapsto \mathrm{e}^x$ on the interval $[0, 1]$, explain graphically, or otherwise, why the polynomial p_1 of degree 1 which minimises $Z_3(f - p_1)$ satisfies the conditions

$$f(0) - p_1(0) = -[f(\tfrac{1}{2}) - p_1](\tfrac{1}{2}) = f(1) - p_1(1)\,.$$

Hence construct this polynomial p_1. Now suppose that $k = 4$, with $t_1 = 0$, $t_2 = \frac{1}{3}$, $t_3 = \frac{2}{3}$, $t_4 = \frac{1}{3}$; use a similar method to construct the polynomial of degree 1 which minimises $Z_4(f - p_1)$.

8.7 Among all polynomials $p_n \in \mathcal{P}_n$ of the form

$$p_n(x) = Ax^n + \sum_{k=0}^{n-1} a_k x^k,$$

where A is a fixed nonzero real number, find the polynomial of best approximation for the function $f(x) \equiv 0$ on the closed interval $[-1, 1]$.

8.8 Find the minimax polynomial $p_n \in \mathcal{P}_n$ on the interval $[-1, 1]$ for the function f defined by

$$f(x) = \sum_{k=0}^{n+1} a_k x^k,$$

where $a_{n+1} \neq 0$.

8.9 Construct the minimax polynomial $p_1 \in \mathcal{P}_1$ on the interval $[-1, 2]$ for the function f defined by $f(x) = |x|$.

8.10 Give a proof of Lemma 8.2.

8.11 Give an example of a continuous real-valued function f defined on the closed interval $[a, b]$ such that the set of critical points for the minimax approximation of f by polynomials from $\mathcal{P}_1$ does not contain either of the points a and b.

8.12 For each nonnegative integer n, and $x \in [0, 1]$, define the Bernstein polynomials $p_{nk} \in \mathcal{P}_n$ by

$$p_{nk}(x) = \frac{n!}{k!(n-k)!} x^k (1-x)^{n-k}, \qquad k = 0, \ldots, n.$$

Show that

$$(1 - x + tx)^n = \sum_{k=0}^{n} p_{nk}(x) t^k;$$

by differentiating this relation successively with respect to t and putting $t = 1$, show that, for any $x \in [0, 1]$,

$$\begin{aligned} \sum_{k=0}^{n} p_{nk}(x) &= 1, \\ \sum_{k=0}^{n} k p_{nk}(x) &= nx, \\ \sum_{k=0}^{n} k(k-1) p_{nk}(x) &= n(n-1)x^2, \end{aligned}$$

and deduce that

$$\sum_{k=0}^{n} (x - k/n)^2 p_{nk}(x) = \frac{x(1-x)}{n}, \qquad x \in [0, 1].$$

Define M to be the upper bound of $|f(x)|$ on $[0, 1]$. Given $\varepsilon > 0$, we can choose $\delta > 0$ such that $|f(x) - f(y)| < \varepsilon/2$ for any x and y in $[0, 1]$ such that $|x - y| < \delta$. Now define the polynomial $p_n \in \mathcal{P}_n$ by

$$p_n(x) = \sum_{k=0}^{n} f(k/n) p_{nk}(x),$$

and choose a fixed value of x in $[0,1]$; show that

$$|f(x) - p_n(x)| \leq \sum_{k=0}^{n} |f(x) - f(k/n)| p_{nk}(x) \,.$$

Using the notation

$$\sum_{k=0}^{n} = \sum_{1} + \sum_{2}$$

where $\sum_1$ denotes the sum over those values of k for which $|x - k/n| < \delta$, and $\sum_2$ denotes the sum over those values of k for which $|x - k/n| \geq \delta$, show that

$$\sum_{1} |f(x) - f(k/n)| p_{nk}(x) < \varepsilon/2 \,.$$

Show also that

$$\sum_{2} |f(x) - f(k/n)| p_{nk}(x) \leq (2M/\delta^2) \sum_{k=0}^{n} (x - k/n)^2 p_{nk}(x) \,.$$

Now, choose $N_0 = M/(\delta^2 \varepsilon)$, and show that

$$|f(x) - p_n(x)| < \varepsilon \qquad \forall\, x \in [0,1] \,,$$

if $n \geq N_0$. Deduce that

$$\|f - p_n\|_\infty < \varepsilon, \quad \text{if } n \geq N_0 \,,$$

where $\|\cdot\|_\infty$ denotes the ∞-norm on the interval $[0,1]$.