

LECTURE 14: *Equioscillation, Part 2*

A direct proof that an optimal minimax approximation $p_* \in \mathcal{P}_n$ must give an equioscillating error is rather tedious, requiring one to chase down the oscillation points one at a time. The following approach is a bit more appealing. We begin with a technical result from which the main theorem will readily follow.

Lemma 2.1. Let $p_* \in \mathcal{P}_n$ be a minimax approximation of $f \in C[a, b]$,

$$\|f - p_*\|_\infty = \min_{p \in \mathcal{P}_n} \|f - p\|_\infty,$$

and let \mathcal{X} denote the set of all points $x \in [a, b]$ for which

$$|f(x) - p_*(x)| = \|f - p_*\|_\infty.$$

Then for all $q \in \mathcal{P}_n$,

$$(2.5) \quad \max_{x \in \mathcal{X}} (f(x) - p_*(x))q(x) \geq 0.$$

Proof. We will prove the lemma by contradiction. Suppose $p_* \in \mathcal{P}_n$ is a minimax approximation, but that (2.5) fails to hold, i.e., there exists some $\tilde{q} \in \mathcal{P}_n$ and $\varepsilon > 0$ such that

$$\max_{x \in \mathcal{X}} (f(x) - p_*(x))\tilde{q}(x) < -2\varepsilon.$$

We first note that $\|\tilde{q}\|_\infty > 0$. Since $(f(x) - p_*(x))\tilde{q}(x)$ is a continuous function on $[a, b]$, it must remain negative on some sufficiently small neighborhood of \mathcal{X} . More concretely, we can find $\delta > 0$ such that

$$(2.6) \quad \max_{x \in \tilde{\mathcal{X}}} (f(x) - p_*(x))\tilde{q}(x) < -\varepsilon,$$

where

$$\tilde{\mathcal{X}} := \{\xi \in [a, b] : \min_{x \in \mathcal{X}} |\xi - x| < \delta\}.$$

To arrive at a contradiction, we will design a function \tilde{p} that better approximates f than p_* , i.e., $\|f - \tilde{p}\|_\infty < \|f - p_*\|_\infty$. This function will take the form

$$\tilde{p}(x) = p_*(x) - \lambda\tilde{q}(x)$$

for (small) constant λ we shall soon determine. Let $E := \|f - p_*\|_\infty$ and pick M such that $|\tilde{q}(x)| \leq M$ for all $x \in \tilde{\mathcal{X}}$. Then for all $x \in \tilde{\mathcal{X}}$,

$$\begin{aligned} |f(x) - \tilde{p}(x)|^2 &= (f(x) - p_*(x))^2 + 2\lambda(f(x) - p_*(x))\tilde{q}(x) + \lambda^2\tilde{q}(x)^2 \\ &= E^2 + 2\lambda(f(x) - p_*(x))\tilde{q}(x) + \lambda^2\tilde{q}(x)^2 \\ (2.7) \quad &< E^2 - 2\lambda\varepsilon + \lambda^2M^2, \end{aligned}$$

This ‘lemma’ is a diluted version of *Kolmogorov’s Theorem*, which is (a) an ‘if and only if’ version of this lemma that (b) appeals to approximation with much more general classes of functions, not just polynomials, and (c) handles complex-valued functions. The proof here is adapted from that more general setting given in Theorem 2.1 of DeVore and Lorentz, *Constructive Approximation* (Springer, 1993).

where this inequality follows from (2.6). To show that \tilde{p} is a better approximation to f than p_* , it will suffice to show that the right-hand side of (2.7) is smaller than E^2 : note that for any $\lambda \in (0, 2\varepsilon/M^2)$, then

$$(2.8) \quad |f(x) - \tilde{p}(x)|^2 < E^2 - 2\lambda\varepsilon + \lambda^2 M^2 < E^2 - \frac{4\varepsilon^2}{M^2} + \frac{4\varepsilon^2}{M^2} = E^2 = \|f - p_*\|^2$$

for all $x \in \tilde{\mathcal{X}}$. Thus \tilde{p} beats p_* on $\tilde{\mathcal{X}}$. Now since \mathcal{X} comprises the points where $|f(x) - p_*(x)|$ attains its maximum, away from $\tilde{\mathcal{X}}$ this error must be bounded away from its maximum, i.e., there exists some $\eta > 0$ such that

$$\max_{\substack{x \in [a,b] \\ x \notin \tilde{\mathcal{X}}}} |f(x) - p_*(x)| \leq E - \eta.$$

Now we want to show that $|f(x) - \tilde{p}(x)| < E$ for these $x \notin \tilde{\mathcal{X}}$ as well. In particular, for such x

$$\begin{aligned} |f(x) - \tilde{p}(x)| &= |f(x) - p_*(x) + \lambda\tilde{q}(x)| \\ &\leq |f(x) - p_*(x)| + \lambda|\tilde{q}(x)| \\ &\leq E - \eta + \lambda\|\tilde{q}\|_\infty, \end{aligned}$$

and so if $\lambda \in (0, \eta/\|\tilde{q}\|_\infty)$,

$$|f(x) - \tilde{p}(x)| < E - \eta + \frac{\eta}{\|\tilde{q}\|_\infty} \|\tilde{q}\|_\infty = E.$$

In conclusion, if

$$\lambda \in \left(0, \min(2\varepsilon/M^2, \eta/\|\tilde{q}\|_\infty)\right),$$

then we constructed $\tilde{p}(x) := p_*(x) - \lambda\tilde{q}(x)$ such that

$$|f(x) - \tilde{p}(x)| < E \quad \text{for all } x \in [a, b],$$

i.e., $\|f - \tilde{p}\|_\infty < \|f - p_*\|$, contradicting the optimality of p_* . ■

With this lemma, we can readily complete the proof of the Oscillation Theorem.

Completion of the Proof of the Oscillation Theorem. We must show that if p_* is a minimax approximation to f , then there exist $n + 2$ points in $[a, b]$ on which the error $f - p_*$ changes sign. If $p_* = f$, the result holds trivially. Suppose then that $\|f - p_*\|_\infty > 0$. In the language of Lemma 2.1, we need to show that (a) the set \mathcal{X} contains (at least) $n + 2$ points and (b) the error oscillates sign at these points. Suppose this is not the case, i.e., we cannot identify $n + 2$ consecutive points in \mathcal{X} at which the error oscillates in sign. Suppose we can only identify

Recall that \mathcal{X} contains all the points $x \in [a, b]$ for which the maximum error is attained: $|f(x) - p_*(x)| = \|f - p_*\|_\infty$.

m such points, $1 \leq m < n + 2$, which we label $x_0 < \dots < x_{m-1}$. We will show how to construct a q that violates Lemma 2.1.

If $m = 1$, $f(x) - p_*(x)$ has the same sign for all $x \in \mathcal{X}$. Set $q(x) = -\text{sgn}(f(x_0) - p_*(x_0))$ (a constant, hence in \mathcal{P}_n), so that $(f(x) - p_*(x))q(x) < 0$ for all $x \in \mathcal{X}$, contradicting Lemma 2.1.

If $m > 1$, the between each consecutive pair of these m points one can then identify $\tilde{x}_1, \dots, \tilde{x}_{m-1}$ where the error changes sign. (See the sketch in the margin.) Then define

$$q(x) = \pm(x - \tilde{x}_1)(x - \tilde{x}_2) \cdots (x - \tilde{x}_{m-1}).$$

Since $m < n + 2$ by assumption, $m - 1 \leq n$, i.e., $q \in \mathcal{P}_n$, so Lemma 2.1 should hold with this choice of q . Since the sign of $q(x)$ does not change between its roots, it does not change within the intervals

$$(a, \tilde{x}_1), (\tilde{x}_1, \tilde{x}_2), \dots, (\tilde{x}_{m-2}, \tilde{x}_{m-1}), (\tilde{x}_{m-1}, b),$$

and the sign of q flips between each of these intervals. Thus the sign of $(f(x) - p_*(x))q(x)$ is the same for all $x \in \mathcal{X}$. Pick the \pm sign in the definition of q such that

$$(f(x) - p_*(x))q(x) < 0 \quad \text{for all } x \in \mathcal{X},$$

thus contradicting Lemma 2.1. Hence, there must be (at least) $n + 2$ consecutive points in \mathcal{X} at which the error flips sign. ■

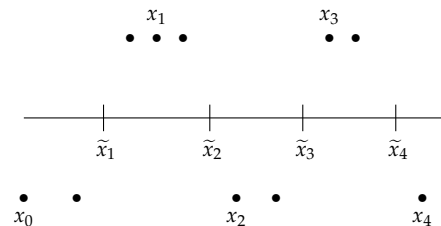
Thus far we have been careful to only speak of *a* minimax approximation, rather than *the* minimax approximation. In fact, the later terminology is more precise, for the minimax approximant is unique.

Theorem 2.3 (Uniqueness of minimax approximant).

The minimax approximant $p_* \in \mathcal{P}_n$ of $f \in C[a, b]$ over the interval $[a, b]$ is unique.

The proof is a straightforward application of the Oscillation Theorem. Suppose p_1 and p_2 are both minimax approximations from \mathcal{P}_n to f on $[a, b]$. Then one can show that $(p_1 + p_2)/2$ is also a minimax approximation. Apply the Oscillation Theorem to obtain $n + 2$ points at which the error for $(p_1 + p_2)/2$ oscillates sign. One can show that these points must also be oscillation points for p_1 and p_2 , and that p_1 and p_2 agree at these $n + 2$ points. Polynomials of degree n that agree at $n + 2$ points must be the same.

This oscillation property forms the basis of algorithms that find the minimax approximation: iteratively adjust an approximating polynomial until it satisfies the oscillation property. The most famous algorithm for computing the minimax approximation is called the



sketch for $m = 5$

• = $f(x) - p_*(x)$ for $x \in \mathcal{X}$

For the full details of this proof, see Theorem 8.5 in Süli and Mayer.

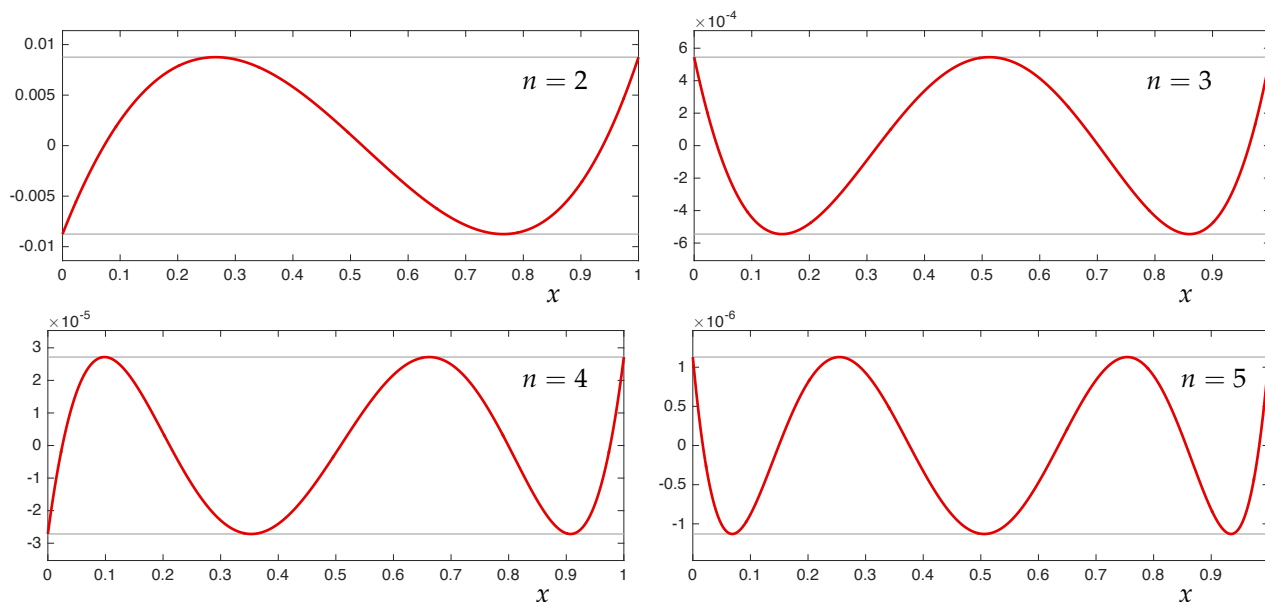


Figure 2.3: Illustration of the equioscillating minimax error $f - p_*$ for approximations of degree $n = 2, 3, 4,$ and 5 with $f(x) = e^x$ for $x \in [a, b]$. In each case, the error attains its maximum with alternating sign at $n + 2$ points.

Remez exchange algorithm, essentially a specialized linear programming procedure. In exact arithmetic, this algorithm is guaranteed to terminate with the correct answer in finitely many operations.

The oscillation property is demonstrated in the Example 2.1, where we approximated $f(x) = e^x$ with a constant. Indeed, the maximum error is attained at two points (that is, $n + 2$, since $n = 0$), and the error differs in sign at those points. Figure 2.3 shows the errors $f(x) - p_*(x)$ for minimax approximations p_* of increasing degree. The oscillation property becomes increasingly apparent as the polynomial degree increases. In each case, there are $n + 2$ extreme points of the error, where n is the degree of the approximating polynomial.

These examples were computed in MATLAB using the Chebfun package's *remez* algorithm. For details, see www.chebfun.org.

Example 2.2 (e^x revisited). Now we shall use the Oscillation Theorem to compute the optimal linear minimax approximation to $f(x) = e^x$ on $[0, 1]$. Assume that the minimax polynomial $p_* \in \mathcal{P}_1$ has the form $p_*(x) = \alpha + \beta x$. Since f is convex, a quick sketch of the situation suggests the maximal error will be attained at the end points of the interval, $x_0 = 0$ and $x_2 = 1$. We assume this to be true, and seek some third point $x_1 \in (0, 1)$ that attains the same maximal error, δ , but with opposite sign. If we can find such a point, then the Oscillation Theorem guarantees that the resulting polynomial is optimal, confirming our assumption that the maximal error was attained at the ends of the interval.

This scenario suggests the following three equations:

$$\begin{aligned} f(x_0) - p_*(x_0) &= \delta \\ f(x_1) - p_*(x_1) &= -\delta \\ f(x_2) - p_*(x_2) &= \delta. \end{aligned}$$

Substituting the values $x_0 = a$, $x_2 = b$, and $p_*(x) = \alpha + \beta x$, these equations become

$$\begin{aligned} 1 - \alpha &= \delta \\ e^{x_1} - \alpha - \beta x_1 &= -\delta \\ e - \alpha - \beta &= \delta. \end{aligned}$$

The first and third equation together imply $\beta = e - 1$. We also deduce that $2\alpha = e^{x_1} - x_1(e - 1) + 1$. A variety of choices for x_1 will satisfy these conditions, but in those cases δ will not be the *maximal error*. We must ensure that

$$|\delta| = \max_{x \in [a, b]} |f(x) - p_*(x)|.$$

To make this happen, require that *the derivative of error* be zero at x_1 , reflecting that the error $f - p_*$ attains a local minimum/maximum at x_1 . (The plots in Figure 2.3 confirm that this is reasonable.) Imposing the condition that $f'(x_1) - p'_*(x_1) = 0$ yields

$$e^{x_1} - \beta = 0.$$

Now we can explicitly solve the equations to obtain

$$\begin{aligned} \alpha &= \frac{1}{2}(e - (e - 1) \log(e - 1)) = 0.89406\dots \\ \beta &= e - 1 = 1.71828\dots \\ x_1 &= \log(e - 1) = 0.54132\dots \\ \delta &= \frac{1}{2}(2 - e + (e - 1) \log(e - 1)) = 0.10593\dots \end{aligned}$$

Figure 2.4 shows the optimal approximation, along with the error $f(x) - p_*(x) = e^x - (\alpha + \beta x)$. In particular, notice the size of the maximum error ($\delta = 0.10593\dots$) and the point $x_1 = 0.54132\dots$ at which this error is attained.

This requirement need not hold at the points x_0 and x_2 , since these points are on the ends of the interval $[a, b]$; it is only required at the interior points where the extreme error is attained, $x_j \in (a, b)$.

Notice that we have a system of four *nonlinear* equations in four unknowns, due to the e^{x_1} term. Generally such nonlinear systems might not have a solution; in this case we can compute one.

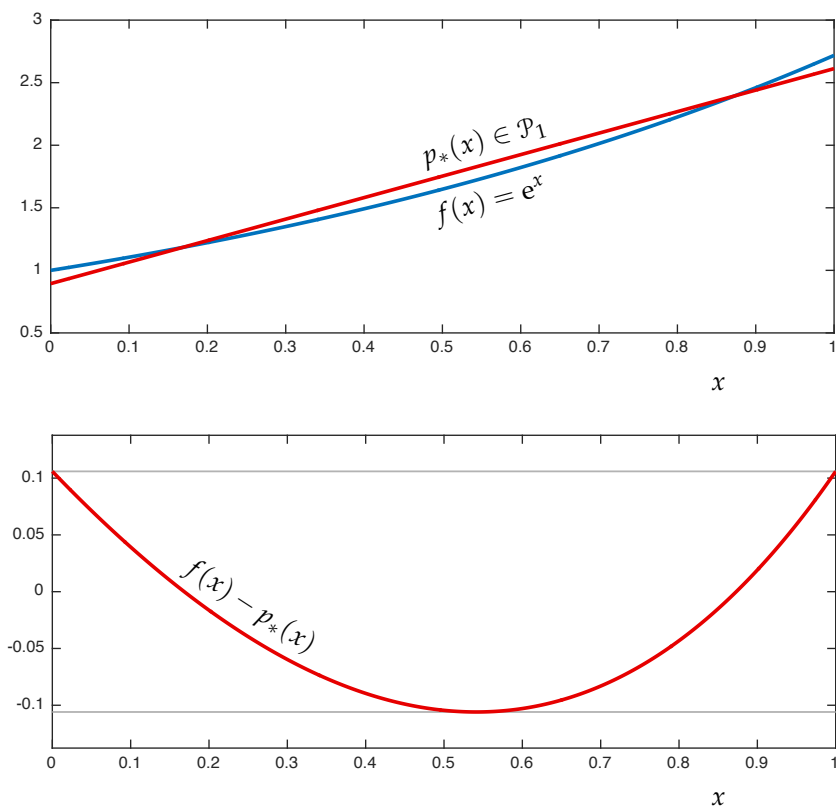


Figure 2.4: The top plot shows the minimax approximation p_* of degree $n = 1$ (red) to $f(x) = e^x$ (blue); the bottom plot shows the error $f(x) - p_*(x)$, equioscillating at $n + 1 = 3$ points.