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 the 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 $q \in \mathcal{P}_n$,

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

This 'lemma' is a diluted version of *Kolmolgorov'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).

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 $\widetilde{q} \in \mathcal{P}_n$ and $\varepsilon > 0$ such that

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

We first note that $\|\widetilde{q}\|_\infty > 0$. Since $\big(f(x) - p_*(x)\big)q(x)$ 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

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

where

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

To arrive at a contradiction, we will design a function that better approximates f than p_*. This function will take the form

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

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

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

where this inequality follows from (2.6). To show that $\widetilde{p}$ 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

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

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

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

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

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

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

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

In conclusion, if

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

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

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

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

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

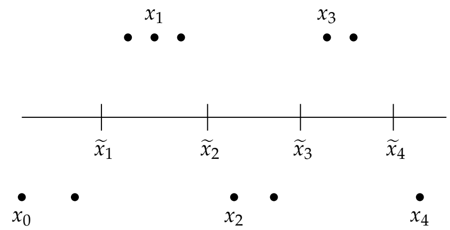
Completion of the Proof of the Oscillation Theorem. We must show that if p_* is a minimax approximation to f, then there exist $n+2$ points in $[a,b]$ on which the error $f - p_*$ changes sign. If $p_* = f$, the result holds trivially. Suppose then that $\|f - p_*\|_\infty > 0$. In the language of Lemma 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 m such points, $1 \le m < n+2$, which we label $x_0 < \cdots < x_{m-1}$.

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

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 $\widetilde{x}_1, \ldots, \widetilde{x}_{m-1}$ where the error changes sign. (See the sketch in the margin.) Then define

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



sketch for $m = 5$
$\bullet = f(x) - p_*(x)$ for $x \in \mathcal{X}$

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 is does not change within the intervals

$$(a, \widetilde{x}_1), \quad (\widetilde{x}_1, \widetilde{x}_2), \quad \cdots, \quad (\widetilde{x}_{m-2}, \widetilde{x}_{m-1}), \quad (\widetilde{x}_{m-1}, b),$$

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

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

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.

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

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 *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.

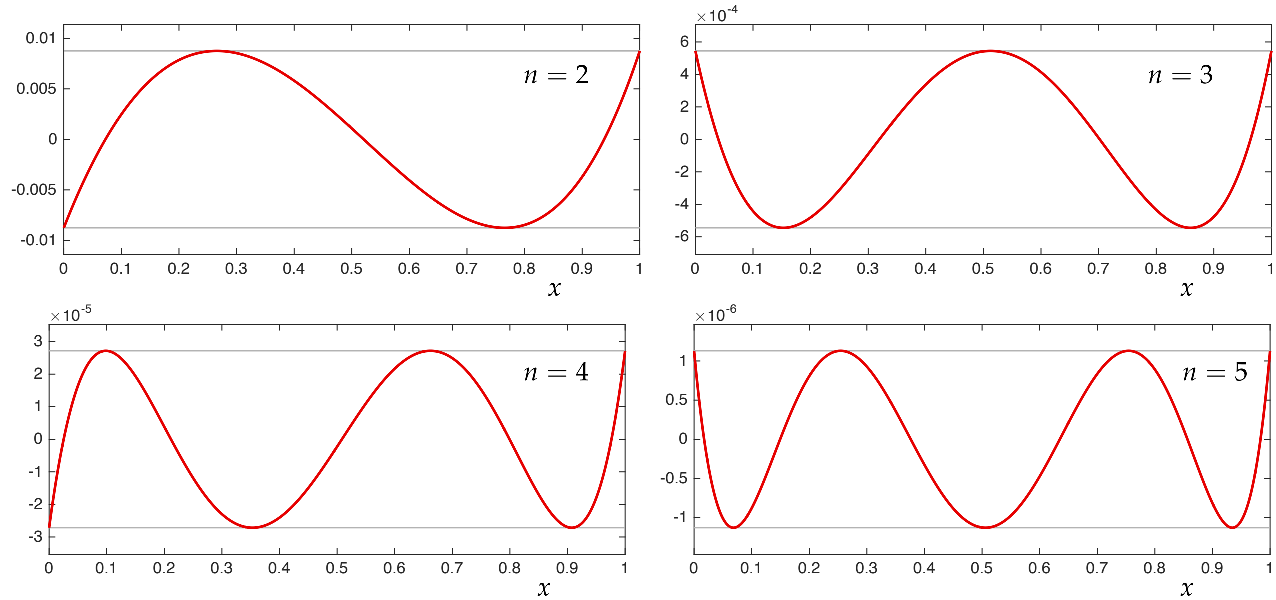


Figure 2.3: Illustration of the equioscillating minimax error $f - p_*$ for approximations of degree $n = 2, 3, 4$, and 5 with $f(x) = \mathrm{e}^x$ for $x \in [a, b]$. In each case, the error attains its maximum with alternating sign at $n + 2$ points.

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

The oscillation property is demonstrated in the Example 2.1, where we approximated $f(x) = \mathrm{e}^x$ 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.

Example 2.2 ($\mathbf{e}^x$ revisited). Now we shall use the Oscillation Theorem to compute the optimal linear minimax approximation to $f(x) = \mathrm{e}^x$ 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

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

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, $x_j \in (a,b)$.

Now we can explicitly solve the equations to obtain

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

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.

The optimal approximation is shown in Figure 2.4; it gives an excellent approximation. The error $f - p_*$ was shown in Figure 2.3.

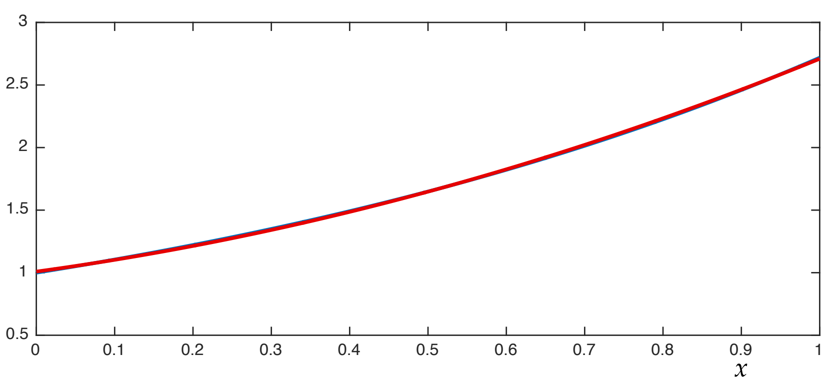


Figure 2.4: Minimax approximation to $f(x) = e^x$ of degree $n = 2$ (red), covering up the function $f(x) = e^x$ (blue, barely visible).

2.3 Optimal Interpolation Points via Chebyshev Polynomials

As an application of the minimax approximation procedure, we consider how best to choose interpolation points $\{x_j\}_{j=0}^n$ to minimize

$$\|f - p_n\|_\infty,$$

where $p_n \in \mathcal{P}_n$ is the interpolant to f at the specified points.

Recall the interpolation error bound developed in Section 1.6: If $f \in C^{n+1}[a,b]$, then for any $x \in [a,b]$ there exists some $\xi \in [a,b]$ such that

$$f(x) - p_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} \prod_{j=0}^{n} (x - x_j).$$

Taking absolute values and maximizing over $[a,b]$ yields the bound

$$\|f - p_n\|_\infty = \max_{\xi \in [a,b]} \frac{|f^{(n+1)}(\xi)|}{(n+1)!} \max_{x \in [a,b]} \Big| \prod_{j=0}^{n} (x - x_j) \Big|.$$

For Runge's example, $f(x) = 1/(1+x^2)$ for $x \in [-5,5]$, we observed that $\|f - p_n\|_\infty \to \infty$ as $n \to \infty$ if the interpolation points $\{x_j\}$ are uniformly spaced over $[-5,5]$. However, Marcinkiewicz's theorem (Section 1.6) guarantees there is always some scheme for assigning the interpolation points such that $\|f - p_n\|_\infty \to 0$ as $n \to \infty$. In the case of Runge's function, we observed that the choice

$$x_j = 5\cos\Big(\frac{j\pi}{n}\Big), \qquad j = 0, \ldots, n$$

is one such scheme. While there is no fail-safe *a priori* system for picking interpolations points that will yield uniform convergence for *all* $f \in C[a,b]$, there is a distinguished choice that works exceptionally well for just about every function you will encounter in practice. We determine this set of interpolation points by choosing those $\{x_j\}_{j=0}^n$ that *minimize the error bound* (which is distinct from – but hopefully akin to – minimizing the error itself, $\|f - p_n\|_\infty$). That is, we want to solve

$$\min_{x_0,\ldots,x_n} \max_{x \in [a,b]} \Big| \prod_{j=0}^{n} (x - x_j) \Big|. \tag{2.9}$$

Notice that

$$\begin{aligned}\prod_{j=0}^{n} (x - x_j) &= x^{n+1} - x^n \sum_{j=0}^{n} x_j + x^{n-1} \sum_{j=0}^{n} \sum_{k=0}^{n} x_j x_k - \cdots + (-1)^{n+1} \prod_{j=0}^{n} x_j \\ &= x^{n+1} - r(x),\end{aligned}$$

where $r \in \mathcal{P}_n$ is a degree-n polynomial depending on the interpolation nodes $\{x_j\}_{j=0}^n$.

To find the optimal interpolation points according to (2.9), we should solve

$$\min_{r\in\mathcal{P}_n}\max_{x\in[a,b]}|x^{n+1}-r(x)| \;=\; \min_{r\in\mathcal{P}_n}\|x^{n+1}-r(x)\|_\infty.$$

Here the goal is to approximate an $(n+1)$-degree polynomial, x^{n+1}, with an n-degree polynomial. The method of solution is somewhat indirect: we will produce a class of polynomials of the form $x^{n+1}-r(x)$ that satisfy the requirements of the Oscillation Theorem, and thus $r(x)$ must be the minimax polynomial approximation to x^{n+1}. As we shall see, the roots of the resulting polynomial $x^{n+1}-r(x)$ will fall in the interval $[a,b]$, and can thus be regarded as 'optimal' interpolation points. For simplicity, we shall focus on the interval $[a,b]=[-1,1]$.

Definition 2.1. The degree-n *Chebyshev polynomial* is defined for $x\in[-1,1]$ by the formula

$$T_n(x)=\cos(n\cos^{-1}x).$$

At first glance, this formula may not appear to define a polynomial at all, since it involves trigonometric functions. But computing the first few examples, we find

Furthermore, it doesn't apply if $|x|>1$. For such x one can define the Chebyshev polynomials using hyperbolic trigonometric functions, $T_n(x)=\cosh(n\cosh^{-1}x)$. Indeed, using hyperbolic trigonometric identities, one can show that this expression generates for $x\notin[-1,1]$ the same polynomials we get for $x\in[-1,1]$ from the standard trigonometric identities.

$$n=0:\quad T_0(x)=\cos(0\cos^{-1}x)=\cos(0)=1$$
$$n=1:\quad T_1(x)=\cos(\cos^{-1}x)=x$$
$$n=2:\quad T_2(x)=\cos(2\cos^{-1}x)=2\cos^2(\cos^{-1}x)-1=2x^2-1.$$

For $n=2$, we employed the identity $\cos 2\theta=2\cos^2\theta-1$, substituting $\theta=\cos^{-1}x$. More generally, we have the identity

$$\cos(n+1)\theta=2\cos\theta\cos n\theta-\cos(n-1)\theta.$$

This formula implies, for $n\ge 2$,

$$T_{n+1}(x)=2xT_n(x)-T_{n-1}(x),$$

a formula related to the three term recurrence used to construct orthogonal polynomials.

In fact, Chebyshev polynomials are orthogonal polynomials on $[-1,1]$ with respect to the inner product

$$\langle f,g\rangle=\int_a^b f(x)g(x)(1-x^2)^{-1/2},$$

a fact we will use when studying Gaussian quadrature later in the semester.

Chebyshev polynomials exhibit a wealth of interesting properties, of which we mention just three.

Theorem 2.4. Let T_n be the degree-n Chebyshev polynomial

$$T_n(x)=\cos(n\cos^{-1}x)$$

for $x\in[-1,1]$.

- $|T_n(x)|\le 1$ for $x\in[-1,1]$.

- The roots of T_n are the n points $\xi_j = \cos\frac{(2j-1)\pi}{2n}$, $j = 1, \ldots, n$.
- For $n \geq 1$, $|T_n(x)|$ is maximized on $[-1, 1]$ at the $n+1$ points $\eta_j = \cos(j\pi/n)$, $j = 0, \ldots, n$:

$$T_n(\eta_j) = (-1)^j.$$

Proof. These results follow from direct calculations. For $x \in [-1, 1]$, $T_n(x) = \cos(n\cos^{-1}(x))$ cannot exceed one in magnitude because cosine cannot exceed one in magnitude. To verify the formula for the roots, compute

$$T_n(\xi_j) = \cos\left(n\cos^{-1}\cos\left(\frac{(2j-1)\pi}{2n}\right)\right) = \cos\left(\frac{(2j-1)\pi}{2}\right) = 0,$$

since cosine is zero at half-integer multiples of π. Similarly,

$$T_n(\eta_j) = \cos\left(n\cos^{-1}\cos\left(\frac{j\pi}{n}\right)\right) = \cos(j\pi) = (-1)^j.$$

Since $T_n(\eta_j)$ is a nonzero degree-n polynomial, it cannot attain more than $n+1$ extrema on $[-1, 1]$, including the endpoint: we have thus characterized all the maxima of $|T_n|$ on $[-1, 1]$. ■

Figure 2.5 shows Chebyshev polynomials T_n for nine different values of n.

2.3.1 *The punchline*

Finally, we are ready to solve the key minimax problem that will reveal optimal interpolation points. Looking at the above plots of Chebyshev polynomials, with their striking equioscillation properties, perhaps you have already guessed the solution yourself.

We defined the Chebyshev polynomials so that

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x)$$

with $T_0(x) = 1$ and $T_1(x) = x$. Thus T_{n+1} has the leading coefficient 2^n for $n \geq 0$. Define

$$\widehat{T}_{n+1} = 2^{-n}T_{n+1}$$

for $n \geq 0$, with $\widehat{T}_0(x) = 1$. These *normalized* Chebyshev polynomials are *monic*, i.e., the leading term in $\widehat{T}_{n+1}(x)$ is x^{n+1}, rather than $2^n x^{n+1}$ as for $T_{n+1}(x)$. Thus, we can write

$$\widehat{T}_{n+1}(x) = x^{n+1} - r_n(x)$$

for some polynomial $r_n(x) = x^{n+1} - \widehat{T}_{n+1}(x) \in \mathcal{P}_n$. We do not especially care about the particular coefficients of this r_n; our quarry will be the *roots* of $\widehat{T}_{n+1}$, the optimal interpolation points.

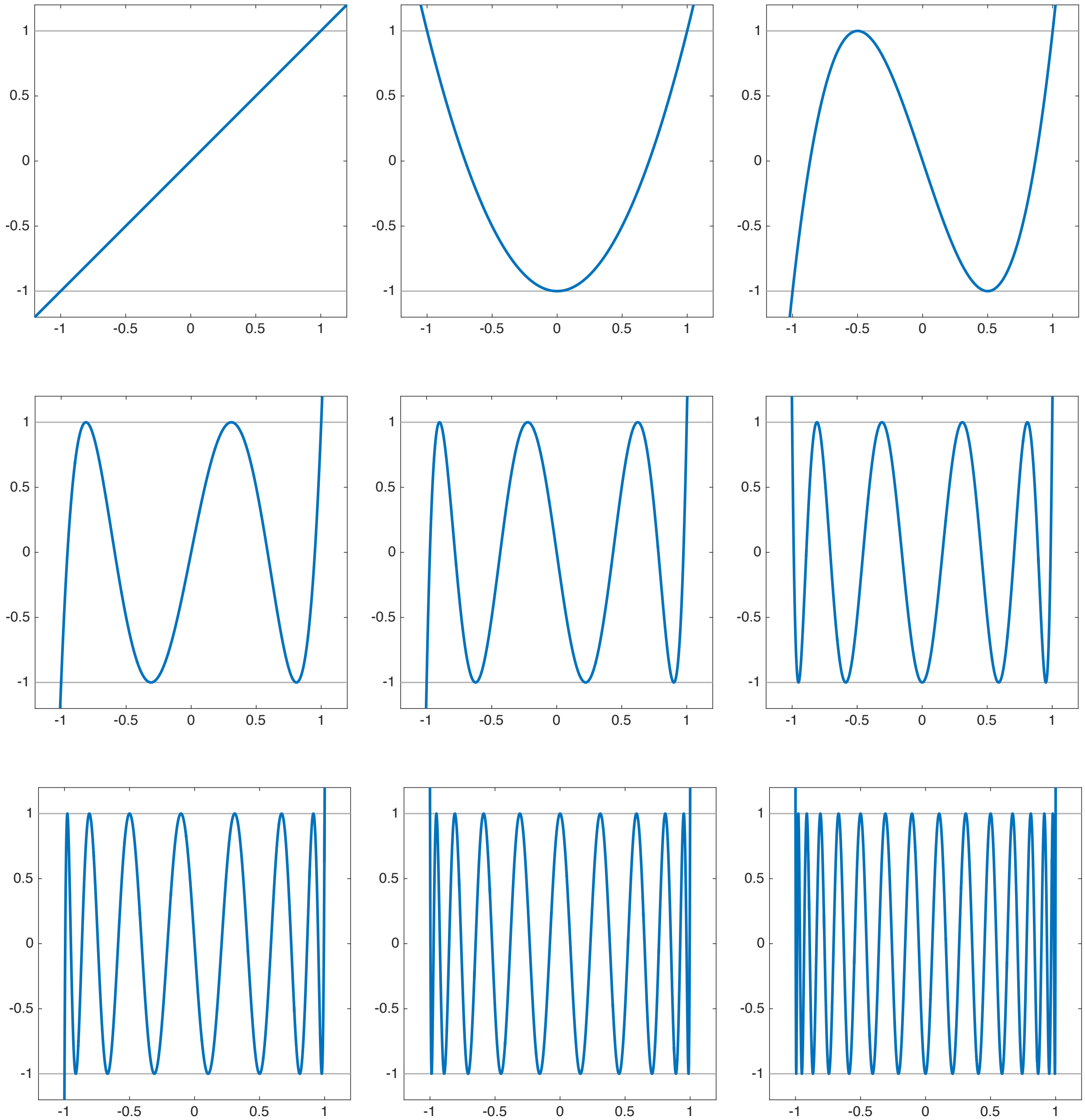


Figure 2.5: Chebyshev polynomials of degree $n = 1, 2, 3$ (top), $n = 5, 7, 10$ (middle), and $n = 15, 20, 30$ (bottom). Note how rapidly these polynomials grow outside the interval $[-1, 1]$.

For $n \geq 0$, the polynomials $\widehat{T}_{n+1}(x)$ oscillate between $\pm 2^{-n}$ for $x \in [-1, 1]$, with the maximal values attained at

$$\eta_j = \cos\left(\frac{j\pi}{n+1}\right)$$

for $j = 0, \ldots, n+1$. In particular,

$$\widehat{T}_{n+1}(\eta_j) = (\eta_j)^{n+1} - r_n(\eta_j) = (-1)^j 2^{-n}.$$

Thus, we have found a polynomial $r_n \in \mathcal{P}_n$, together with $n+2$ distinct points, $\eta_j \in [-1, 1]$ where the maximum error

$$\max_{x \in [-1,1]} |x^{n+1} - r_n(x)| = 2^{-n}$$

is attained with alternating sign. Thus, by the oscillation theorem, we have found the minimax approximation to x^{n+1}.

Theorem 2.5 (Optimal approximation of x^{n+1})**.**
The optimal approximation to x^{n+1} from $\mathcal{P}_n$ for $x \in [-1, 1]$ is

$$r_n(x) = x^{n+1} - \widehat{T}_{k+1}(x) = x^{n+1} - 2^{-n} T_{k+1}(x) \in \mathcal{P}_n.$$

Thus, the optimal interpolation points are those $n+1$ roots of $x^{n+1} - r_n(x)$, that is, the roots of the degree-$(n+1)$ Chebyshev polynomial:

$$\xi_j = \cos\left(\frac{(2j+1)\pi}{2n+2}\right), \quad j = 0, \ldots, n.$$

For generic intervals $[a, b]$, a change of variable demonstrates that the same points, appropriately shifted and scaled, will be optimal.

Similar properties hold if interpolation is performed at the $n+1$ points

$$\eta_j = \cos\left(\frac{j\pi}{n}\right), \quad j = 0, \ldots, n,$$

which are also called Chebyshev points and are perhaps more popular due to their slightly simpler formula. (We used these points to successfully interpolate Runge's function, scaled to the interval $[-5, 5]$.) While these points differ from the roots of the Chebyshev polynomial, they *have the same distribution* as $n \to \infty$. That is the key.

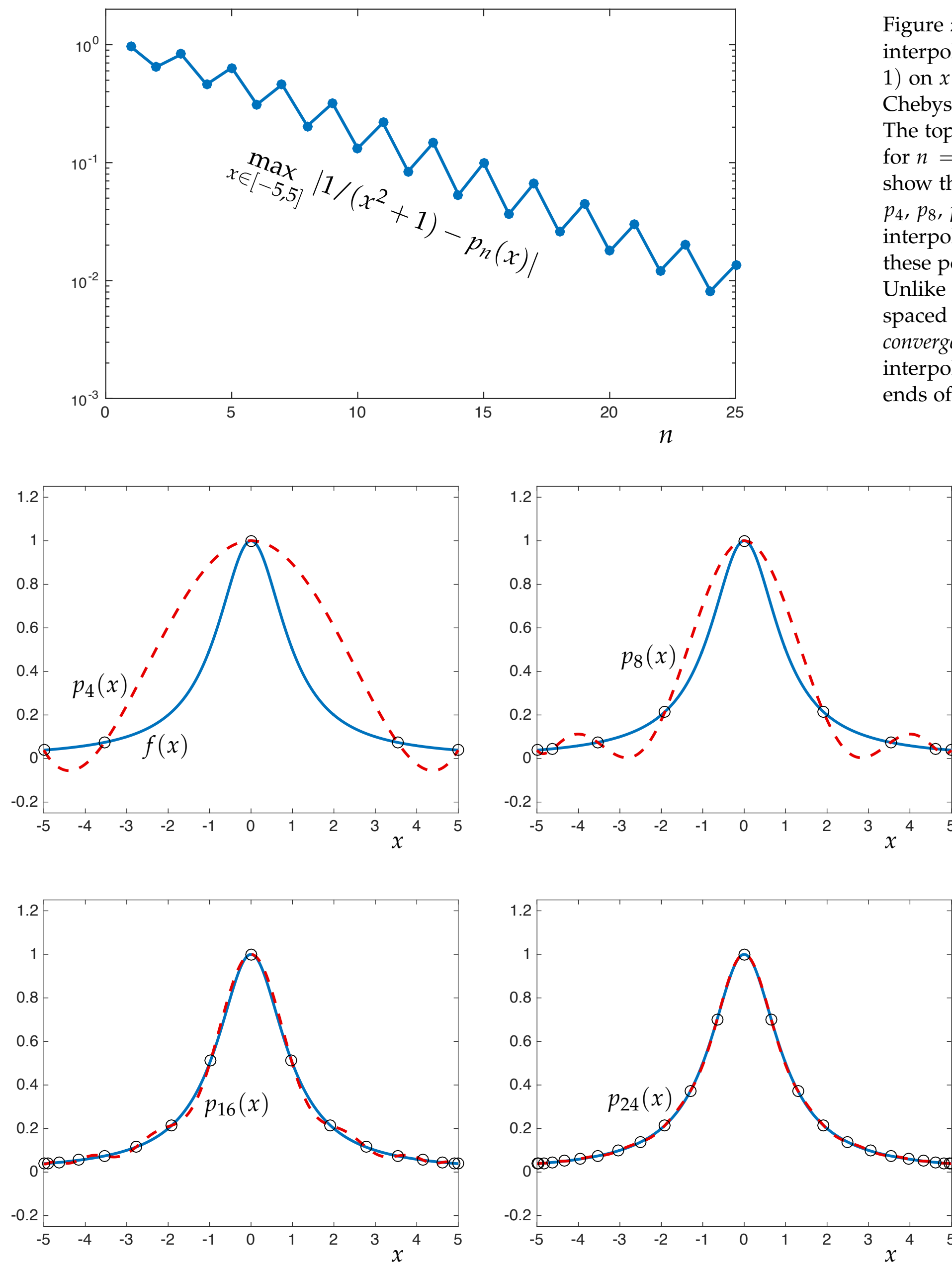


Figure 2.6: Repetition of Figure 1.8, interpolating Runge's function $1/(x^2+1)$ on $x \in [-5,5]$, but now using Chebyshev points $x_j = 5\cos(j\pi/n)$. The top plot shows this convergence for $n = 0,\ldots,25$; the bottom plots show the interpolating polynomials p_4, p_8, p_{16}, and p_{24}, along with the interpolation points that determine these polynomials (black circles). Unlike interpolation at uniformly spaced points, these interpolants *do converge* to f as $n \to \infty$. Notice how the interpolation points cluster toward the ends of the domain $[-5,5]$.