Chebyshev Interpolation

Notes by Linus Tang.

These notes have not been thoroughly reviewed. Any errors below are my own responsibility.

Sources:
- Wikipedia editors
  - https://en.wikipedia.org/wiki/Chebyshev_nodes
  - https://en.wikipedia.org/wiki/Chebyshev_polynomials
- Folklore

Assumed background knowledge:
- Familiarity with polynomial interpolation.
- A bit of real analysis. Good intuition will be more important than formal background.

Contents

Main problem: estimating functions with low-degree polynomials

Consider the problem of estimating a function $f : [x_{\min}, x_{\max}] \to \mathbb{R}$ with a polynomial P_n of degree at most $n - 1$. Say we want to minimize the "L_∞ error", which is

$$\max_{x \in [x_{\min}, x_{\max}]} |f(x) - P_n(x)|.$$

In order for a good degree $n - 1$ approximation to exist, it turns out that $f^{(n)}(\cdot)$ (denoting the n-th derivative of f) should be small. Why? Because your error term $f_e = f - P_n$ has the same nth derivative as f, and you want f_e to be close to 0 on an interval. Hopefully this line of reasoning should feel intuitive, but if it doesn't, the rest of these notes should help you build this intuition a bit.

Interpolation

Theorem 1. Let I be the interval $[x_{\min}, x_{\max}]$ and $f : I \to \mathbb{R}$ with continuous n-th derivative. Furthermore, suppose $\left|f^{(n)}(x)\right| \leq M$ for all $x \in I$. Let $x_0, ..., x_{n-1}$ be n distinct points on I. Choose P_n to be the unique polynomial of degree at most $n - 1$ such that $P_n(x_k) = f(x_k)$ for $k = 0, ..., n - 1$. Then for all $x \in I$, we have

$$|f(x) - P_n(x)| \leq \left|\frac{M}{n!}(x - x_0) \cdots (x - x_{n-1})\right|.$$

Proof of Theorem 1. Let $f_e = f - P_n$ be the error term. Since P_n is a polynomial of degree at most $n - 1$, it follows that $P_n^{(n)} = 0$, so $f_e^{(n)} = f^{(n)}$. Furthermore, $f^{(n)}$ vanishes on $x_0, ..., x_{n-1}$ by definition.

We have reduced the problem to the following: Suppose $|f_e^{(n)}|$ is bounded on I by M and f_e vanishes on $x_0, ..., x_{n-1}$. Then for all $x \in I$, we want to show

$$|f_e(x)| \leq \left|\frac{M}{n!}(x - x_0) \cdots (x - x_{n-1})\right|.$$

In particular, we've reduced the problem to something not involving the original function f, only the error term f_e.

Focus on arbitrary fixed $x \in I$. Let P_{n+1} be a polynomial of degree at most n that, like f_e, vanishes on $x_0, ..., x_{n-1}$, and furthermore satisfies $P_{n+1}(x) = f_e(x) = y$. In particular, P_{n+1} agrees with f_e on at least $n + 1$ points.

We can compute the polynomial P_{n+1} by Lagrange interpolation; it is given by

$$P_{n+1}(u) = y \cdot \frac{(u - x_0) \cdots (u - x_{n-1})}{(x - x_0) \cdots (x - x_{n-1})}.$$

Since P_{n+1} has degree at most n, the n-th derivative of P_{n+1} is constant, and we can see it is given by

$$P_{n+1}^{(n)}(u) = \frac{y \cdot n!}{(x - x_0) \cdots (x - x_{n-1})}.$$

Now it suffices to show that $M \geq \left|f_e^{(n)}(u)\right| \geq \left|P_{n+1}^{(n)}(u)\right|$ for some u.

We can reduce this to a problem about $f_e - P_{n+1}$. In particular, we finish by applying the following proposition to $f_e - P_{n+1}$:

Proposition. Let $g : I \to \mathbb{R}$ have continuous n-th derivative and vanish at $n+1$ distinct points. Then $g^{(n)}(u_+) \geq 0$ for some $u \in I$ and $g^{(n)}(u_-) \leq 0$ for some $u_- \in I$.

The proof of this proposition is by induction on n, using the fact that if a (continuously n-th differentiable) vanishes at $\geq n+1$ points, then its first derivative vanishes at $\geq n$ points.

To wrap up, for all $x \in I$, if $x \neq x_0, ..., x_{n-1}$, then $f_e - P_{n+1}$ vanishes at $n+1$ distinct points $x, x_0, ..., x_{n-1}$. then applying the above proposition tells us that there exists u such that

$$M \geq \left|f_e^{(n)}(u)\right| \geq \left|P_{n+1}^{(n)}(u)\right| = \left|\frac{y \cdot n!}{(x - x_0) \cdots (x - x_{n-1})}\right|,$$

and it follows that

$$|f_e(x)| = |y| \leq \left|\frac{M}{n!}(x - x_0) \cdots (x - x_{n-1})\right|,$$

proving the theorem. □

Aside: repeated x_k and Talor series

It turns out that Taylor series are a specific case of the above idea. Suppose we don't require $x_0, ..., x_{n-1}$ to be distinct. Instead, if x_k appears ℓ times in $x_0, ..., x_{n-1}$, then we require f and its first $\ell - 1$ derivatives to agree with P_n and its first $\ell - 1$ derivatives at x_k. Formally,

$$f^{(j)}(x_k) = P_n^{(j)}(x_k)$$

for all $j < \ell$. The polynomial P_n of degree at most $n-1$ which satisfies these constraints is still existent and unique. With some care, we can show that Theorem 1 generalizes accordingly. (Replacing equality constraints with derivative equality constraints to deal with multiplicity is a common theme in real analysis.)

Now, the degree $n-1$ Taylor series of f at t just comes from setting $x_0 = \cdots = x_{n-1} = t$, and taking the polynomial P_n.

Chebyshev Interpolation

If we want to use Theorem 1 to get an approximation with a guaranteed low L_∞ error, naturally we want to interpolate at points $x_0, ..., x_{n-1}$ that minimize

$$\max_{x \in I}(x - x_0) \cdots (x - x_{n-1}).$$

(So, we can see that while setting $x_0 = \cdots = x_{n-1} = t$ as in the case of Taylor series is a good choice in the setting $x \approx t$, it is bad at minimizing the worst-case error over a large interval, since $(x - t)^n$ can be large).

It turns out that when $I = [-1, 1]$, the best choice of $x_0, ..., x_{n-1}$ is given by what are called Chebyshev nodes of the first kind,

$$x_k = \cos\left(\frac{\left(k + \frac{1}{2}\right)\pi}{n}\right).$$

Formally,

Theorem 2. Let n be a positive integer and $I = [-1, 1]$. Then

$$\min_{(x_0^*,\ldots,x_{n-1}^*)\in\mathbb{R}^n} \max_{x\in I}(x-x_0^*)\cdots(x-x_{n-1}^*) = 2^{1-n}$$

with the minimum achieved by $x_k^* = \cos\left(\frac{(k+\frac{1}{2})\pi}{n}\right)$.

In order to prove this, we first establish the existence of a series of polynomials called the Chebyshev polynomials of the first kind. The relevant properties are summarized by the following theorem.

Theorem 3. (Chebyshev polynomials of the first kind) For any nonnegative integer n, there exists a degree n polynomial T_n such that

$$T_n(\cos\theta) = \cos(n\theta)$$

for all θ. Furthermore, when n is positive, the x^n coefficient is 2^{n-1}.

Proof of Theorem 3. We can write the condition as $T_n\left(\frac{e^{i\theta}+e^{-i\theta}}{2}\right) = T_n\left(\frac{e^{in\theta}+e^{-in\theta}}{2}\right)$, so it suffices to have

$$T_n\left(\frac{s+\frac{1}{s}}{2}\right) = \left(\frac{s^n+\frac{1}{s^n}}{2}\right).$$

Now $T_0(x) = 1$ and $T_1(x) = x$ satisfy the above, and the polynomials given by the recursion

$$T_j = 2xT_{j-1}(x) - T_{j-2}(x)$$

work due to the identity

$$s^j + \frac{1}{s^j} = \left(s+\frac{1}{s}\right)\left(s^{j-1}+\frac{1}{s^{j+1}}\right) - \left(s^{j-2}+\frac{1}{s^{j-2}}\right).$$

The claim about the leading coefficent follows from the recursion. □

Proof of Theorem 2. Define polynomial T_n as in Theorem 3, so T_n has a root at $x_k = \cos\left(\frac{(k+\frac{1}{2})\pi}{n}\right)$ for $k = 0, ..., n-1$. Since T_n has a leading coefficient of 2^{n-1}, we have

$$(x-x_0)\cdots(x-x_{n-1}) = 2^{1-n}\cdot T_n(x).$$

Then we have

$$\max_{x\in I}(x-x_0^*)\cdots(x-x_{n-1}^*) = 2^{1-n}\cdot \max_{x\in I} T_n(x) = 2^{1-n},$$

proving the upper bound.

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

Let $(x_0^*, ..., x_{n-1}^*) \in \mathbb{R}^n$ and $P(x) = (x-x_0^*)\cdots(x-x_{n-1}^*)$.

If we have $\max_{x\in I} P(x) < 2^{1-n}$, then we consider $Q(x) = P(x) - 2^{1-n}\cdot T_n(x)$.

In particular, since Q is a difference of monic degree n polynomials, it has degree at most $n-1$. However, we have that $\left(Q(\cos 0), Q\left(\cos\frac{\pi}{n}\right), Q\left(\cos\frac{2\pi}{n}\right), ..., Q(\cos\pi)\right)$ alternates between positive and negative, implying that Q has at least n roots, contradiction.

Therefore, $\max_{x\in I} P(x) \geq 2^{1-n}$ for all $(x_0^*, ..., x_{n-1}^*) \in \mathbb{R}^n$, where $P(x) = (x-x_0^*)\cdots(x-x_{n-1}^*)$, proving the lower bound. □

For completeness, we substitute these "optimal" values of $x_0, ..., x_{k-1}$ into Theorem 1.

Theorem 4. (Chebyshev interpolant) Let I be the interval $[x_{\min}, x_{\max}]$ and $f : I \to \mathbb{R}$ with continuous n-th derivative. Furthermore, suppose $\left|f^{(n)}(x)\right| \leq M$ for all $x \in I$.

Define $x_k = \frac{x_{\min}+x_{\max}}{2} + \frac{x_{\max}-x_{\min}}{2} \cdot \cos\left(\frac{(k+\frac{1}{2})\pi}{n}\right)$.

Choose P_n to be the unique polynomial of degree at most $n-1$ such that $P_n(x_k) = f(x_k)$ for $k = 0, ..., n-1$. Then for all $x \in I$, we have

$$|f(x) - P_n(x)| \leq 2 \cdot \frac{M}{n!} \cdot \left(\frac{x_{\max} - x_{\min}}{4}\right)^n.$$

Proof sketch of Theorem 4. Generalize the result of Theorem 2 by linearly transforming the interval $[-1, 1]$ to $[x_{\min}, x_{\max}]$. Then substitute the resulting values of $x_0, ..., x_{n-1}$ into Theorem 1. □

In this case, we call P_n the n-th Chebyshev interpolant of f.

Here's an example which shows the power of Chebyshev interpolants.

Example. Suppose you want a degree ≤ 15 polynomial which estimates $\sin(x)$ over $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$. If you use Chebyshev interpolation, your polynomial P_{16} satisfies

$$|\sin(x) - P_{16}(x)| \leq \frac{2 \cdot \left(\frac{\pi}{4}\right)^{16}}{16!} \approx 2 \times 10^{-15}$$

for all $x \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$.

Aside: computing coefficients

One nice property of Chebyshev interpolation is that its coefficients in a certain basis can be computed by a discrete Fourier transform. Specifically:

Since $T_0, ..., T_{n-1}$ is a sequence of polynomials with degrees $0, ..., n-1$, respectively, and P_n is a polynomial of degree $n-1$, we can write

$$P_n(x) = \sum_{j=0}^{n-1} c_j T_j(x)$$

for some coefficients $c_0, ..., c_{n-1}$.

Recalling that $T_j(\cos(\theta)) = \cos(j\theta)$ and $x_k = \cos\left(\frac{(k+\frac{1}{2})\pi}{n}\right)$, it follows that

$$M \begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_{n-1} \end{pmatrix} = \begin{pmatrix} f(x_0) \\ f(x_1) \\ \vdots \\ f(x_{n-1}) \end{pmatrix},$$

where M is a matrix with entries $M_{jk} = \cos\left(\frac{j(k+\frac{1}{2})\pi}{n}\right)$.

Therefore, we can compute the coefficients $c_0, ..., c_{n-1}$ by multiplying the vector of evaluations of f by M^{-1}. This is a discrete Fourier transform!

Summary

- You can get a degree $\leq n-1$ polynomial estimate P_n of a function $f : I \to \mathbb{R}$ by interpolating f at n points.
- This idea generalizes the Taylor series, which takes all n points to be the same point.
- The resulting bound on the worst-case error can be minimized by interpolating at the specific points

$$x_k = \cos\left(\frac{\left(k + \frac{1}{2}\right)\pi}{n}\right).$$

- These points are roots of the Chebyshev polynomials of the first kind.
- The Chebyshev polynomials are a basis over which the coefficients of P_n are a discrete Fourier transform of the relevant evaluations of f.