Tomorrow is the course registration deadline

I will send out invitations to the Piazza message board for anyone who has registered for the course

Course announcements will be posted on Piazza

If you are auditing but would like to receive announcements, please email chr@seas.harvard.edu

Homework 1 is now available

My office hours today: Pierce Hall 305, 1pm–3pm

Full office hour schedule is posted

TF Yue is organizing a mini Python tutorial on Wednesday afternoon—see Piazza for details
Last time

- Last time: gave motivation for data fitting
- Today: polynomial interpolation for:
  1. A discrete set of points
  2. Continuous functions
We would like to avoid these kinds of sensitivities to perturbations . . . How can we do better?

Try to construct a basis such that the interpolation matrix is the identity matrix

This gives a condition number of 1, and as an added bonus we also avoid inverting a dense $(n + 1) \times (n + 1)$ matrix
Lagrange Interpolation

Key idea: Construct basis \( \{ L_k \in \mathbb{P}_n, k = 0, \ldots, n \} \) such that

\[
L_k(x_i) = \begin{cases} 
0, & i \neq k, \\
1, & i = k.
\end{cases}
\]

The polynomials that achieve this are called Lagrange polynomials\(^1\)

See Lecture: These polynomials are given by:

\[
L_k(x) = \prod_{j=0, j \neq k}^{n} \frac{x - x_j}{x_k - x_j}
\]

and then the interpolant can be expressed as

\[
p_n(x) = \sum_{k=0}^{n} y_k L_k(x)
\]

\(^1\)Joseph-Louis Lagrange, 1736–1813
Lagrange Interpolation

Two Lagrange polynomials of degree 5
Hence we can use Lagrange polynomials to interpolate discrete data (recall plot from I.1)

We have essentially solved the problem of interpolating discrete data perfectly!

With Lagrange polynomials we can construct an interpolant of discrete data with condition number of 1
Interpolation for Function Approximation
Interpolation for Function Approximation

We now turn to a different (and much deeper) question: Can we use interpolation to accurately approximate continuous functions?

Suppose the interpolation data come from samples of a continuous function $f$ on $[a, b] \subset \mathbb{R}$

Then we’d like the interpolant to be “close to” $f$ on $[a, b]$

The error in this type of approximation can be quantified from the following theorem due to Cauchy$^2$:

$$f(x) - p_n(x) = \frac{f^{(n+1)}(\theta)}{(n+1)!} (x-x_0) \cdots (x-x_n) \text{ for some } \theta \in (a, b)$$

---

$^2$Augustin-Louis Cauchy, 1789–1857
Polynomial Interpolation Error

We prove this result in the case $n = 1$

Let $p_1 \in \mathbb{P}_1[x_0, x_1]$ interpolate $f \in C^2[a, b]$ at $\{x_0, x_1\}$

For some $\lambda \in \mathbb{R}$, let

$$q(x) \equiv p_1(x) + \lambda(x - x_0)(x - x_1),$$

here $q$ is quadratic and interpolates $f$ at $\{x_0, x_1\}$

Fix an arbitrary point $\hat{x} \in (x_0, x_1)$ and set $q(\hat{x}) = f(\hat{x})$ to get

$$\lambda = \frac{f(\hat{x}) - p_1(\hat{x})}{(\hat{x} - x_0)(\hat{x} - x_1)}$$

Goal: Get an expression for $\lambda$, since then we obtain an expression for $f(\hat{x}) - p_1(\hat{x})$
Polynomial Interpolation Error

Now, let \( e(x) \equiv f(x) - q(x) \)

- \( e \) has 3 roots in \([x_0, x_1]\), i.e. at \( x = x_0, \hat{x}, x_1 \)
- Therefore \( e' \) has 2 roots in \((x_0, x_1)\) (by Rolle’s theorem)
- Therefore \( e'' \) has 1 root in \((x_0, x_1)\) (by Rolle’s theorem)

Let \( \theta \in (x_0, x_1) \) be such that\(^3\) \( e''(\theta) = 0 \)

Then

\[
0 = e''(\theta) = f''(\theta) - q''(\theta)
\]

\[
= f''(\theta) - p''_1(\theta) - \lambda \frac{d^2}{d\theta^2} (\theta - x_0)(\theta - x_1)
\]

\[
= f''(\theta) - 2\lambda
\]

hence \( \lambda = \frac{1}{2} f''(\theta) \)

\(^3\)Note that \( \theta \) is a function of \( \hat{x} \)
Polynomial Interpolation Error

Hence, we get

\[ f(\hat{x}) - p_1(\hat{x}) = \lambda(\hat{x} - x_0)(\hat{x} - x_1) = \frac{1}{2}f''(\theta)(\hat{x} - x_0)(\hat{x} - x_1) \]

for any \( \hat{x} \in (x_0, x_1) \) (recall that \( \hat{x} \) was chosen arbitrarily)

This argument can be generalized to \( n > 1 \) to give

\[ f(x) - p_n(x) = \frac{f^{(n+1)}(\theta)}{(n+1)!}(x-x_0) \cdots (x-x_n) \text{ for some } \theta \in (a, b) \]
Polynomial Interpolation Error

For any $x \in [a, b]$, this theorem gives us the error bound

$$|f(x) - p_n(x)| \leq \frac{M_{n+1}}{(n+1)!} \max_{x \in [a,b]} |(x - x_0) \ldots (x - x_n)|,$$

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

If $1/(n+1)! \to 0$ faster than

$$M_{n+1} \max_{x \in [a,b]} |(x - x_0) \ldots (x - x_n)| \to \infty$$

then $p_n \to f$

Unfortunately, this is not always the case!
Runge’s Phenomenon

A famous pathological example of the difficulty of interpolation at equally spaced points is Runge’s Phenomenon

Consider $f(x) = 1/(1 + 25x^2)$ for $x \in [-1, 1]$
Runge’s Phenomenon

Note that of course $p_n$ fits the evenly spaced samples exactly.

But we are now also interested in the maximum error between $f$ and its polynomial interpolant $p_n$.

That is, we want $\max_{x \in [-1,1]} |f(x) - p_n(x)|$ to be small!

This is generally referred to as the “infinity norm” or the “max norm”:

$$\|f - p_n\|_\infty \equiv \max_{x \in [-1,1]} |f(x) - p_n(x)|$$
Interpolating Runge’s function at evenly spaced points leads to exponential growth of infinity norm error!

We would like to construct an interpolant of $f$ such that this kind of pathological behavior is impossible.
Minimizing Interpolation Error

To do this, we recall our error equation

\[ f(x) - p_n(x) = \frac{f^{n+1}(\theta)}{(n + 1)!}(x - x_0) \ldots (x - x_n) \]

We focus our attention on the polynomial \((x - x_0) \ldots (x - x_n)\), since we can choose the interpolation points

Intuitively, we should choose \(x_0, x_1, \ldots, x_n\) such that \(\|(x - x_0) \ldots (x - x_n)\|_\infty\) is as small as possible
Interpolation at Chebyshev Points

Result from Approximation Theory:
For $x \in [-1, 1]$, the minimum value of $\| (x - x_0) \ldots (x - x_n) \|_\infty$ is $1/2^n$, achieved by the polynomial $T_{n+1}(x)/2^n$

$T_{n+1}(x)$ is the Chebyshev poly. (of the first kind) of order $n + 1$ ($T_{n+1}$ has leading coefficient of $2^n$, hence $T_{n+1}(x)/2^n$ is monic)

Chebyshev polys “equi-oscillate” between $-1$ and $1$, hence it’s not surprising that they are related to the minimum infinity norm
Interpolation at Chebyshev Points

Chebyshev polynomials are defined for $x \in [-1, 1]$ by
$$T_n(x) = \cos(n \cos^{-1} x), \quad n = 0, 1, 2, \ldots$$

Or equivalently\(^4\), the recurrence relation,
$$T_0(x) = 1,$$
$$T_1(x) = x,$$
$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x), \quad n = 1, 2, 3, \ldots$$

To set $(x - x_0) \ldots (x - x_n) = T_{n+1}(x)/2^n$, we choose interpolation points to be the roots of $T_{n+1}$

**Exercise:** Show that the roots of $T_n$ are given by
$$x_j = \cos((2j - 1)\pi/2n), \quad j = 1, \ldots, n$$

\(^4\)Equivalence can be shown using trig. identities for $T_{n+1}$ and $T_{n-1}$
Interpolation at Chebyshev Points

We can combine these results to derive an error bound for interpolation at “Chebyshev points”

Generally speaking, with Chebyshev interpolation, \( p_n \) converges to any smooth \( f \) very rapidly! e.g. Runge function:

If we want to interpolate on an arbitrary interval, we can map Chebyshev points from \([-1, 1]\) to \([a, b]\)
Interpolation at Chebyshev Points

Note that convergence rates depend on smoothness of $f$—precise statements about this can be made, outside the scope of AM205.

In general, smoother $f \implies$ faster convergence$^5$

e.g. [ch_inter.py] compare convergence of Chebyshev interpolation of Runge’s function (smooth) and $f(x) = |x|$ (not smooth)

$^5$For example, if $f$ is analytic, we get exponential convergence!
Another View on Interpolation Accuracy

We have seen that the interpolation points we choose have an enormous effect on how well our interpolant approximates $f$

The choice of Chebyshev interpolation points was motivated by our interpolation error formula for $f(x) - p_n(x)$

But this formula depends on $f$ — we would prefer to have a measure of interpolation accuracy that is independent of $f$

This would provide a more general way to compare the quality of interpolation points . . . This is provided by the Lebesgue constant
Let $\mathcal{X}$ denote a set of interpolation points, 
\[ \mathcal{X} \equiv \{ x_0, x_1, \ldots, x_n \} \subset [a, b] \]

A fundamental property of $\mathcal{X}$ is its Lebesgue constant, $\Lambda_n(\mathcal{X})$,
\[
\Lambda_n(\mathcal{X}) = \max_{x \in [a, b]} \sum_{k=0}^{n} |L_k(x)|
\]

The $L_k \in \mathbb{P}_n$ are the Lagrange polynomials associated with $\mathcal{X}$, hence $\Lambda_n$ is also a function of $\mathcal{X}$

$\Lambda_n(\mathcal{X}) \geq 1$, why?
Lebesgue Constant

Think of polynomial interpolation as a map, $\mathcal{I}_n$, where

$\mathcal{I}_n : C[a, b] \rightarrow \mathbb{P}_n[a, b]$

$\mathcal{I}_n(f)$ is the degree $n$ polynomial interpolant of $f \in C[a, b]$ at the interpolation points $\mathcal{X}$

**Exercise:** Convince yourself that $\mathcal{I}_n$ is linear (e.g. use the Lagrange interpolation formula)

The reason that the Lebesgue constant is interesting is because it bounds the “operator norm” of $\mathcal{I}_n$:

$$
\sup_{f \in C[a, b]} \frac{\|\mathcal{I}_n(f)\|_\infty}{\|f\|_\infty} \leq \Lambda_n(\mathcal{X})
$$
Lebesgue Constant

Proof:

\[ \|I_n(f)\|_{\infty} = \left\| \sum_{k=0}^{n} f(x_k) L_k \right\|_{\infty} = \max_{x \in [a,b]} \left| \sum_{k=0}^{n} f(x_k) L_k(x) \right| \]

\[ \leq \max_{x \in [a,b]} \sum_{k=0}^{n} |f(x_k)| \|L_k(x)\| \]

\[ \leq \left( \max_{k=0,1, \ldots, n} |f(x_k)| \right) \max_{x \in [a,b]} \sum_{k=0}^{n} |L_k(x)| \]

\[ \leq \|f\|_{\infty} \max_{x \in [a,b]} \sum_{k=0}^{n} |L_k(x)| \]

\[ = \|f\|_{\infty} \Lambda_n(\mathcal{X}) \]

Hence

\[ \frac{\|I_n(f)\|_{\infty}}{\|f\|_{\infty}} \leq \Lambda_n(\mathcal{X}), \quad \text{so} \quad \sup_{f \in C[a,b]} \frac{\|I_n(f)\|_{\infty}}{\|f\|_{\infty}} \leq \Lambda_n(\mathcal{X}). \]
Lebesgue Constant

The Lebesgue constant allows us to bound interpolation error in terms of the smallest possible error from $\mathbb{P}_n$

Let $p_n^* \in \mathbb{P}_n$ denote the best infinity-norm approximation to $f$, i.e.
$$\|f - p_n^*\|_{\infty} \leq \|f - w\|_{\infty} \text{ for all } w \in \mathbb{P}_n$$

Some facts about $p_n^*$:
- $\|p_n^* - f\|_{\infty} \to 0$ as $n \to \infty$ for any continuous $f$! (Weierstraß approximation theorem)
- $p_n^* \in \mathbb{P}_n$ is unique
- In general, $p_n^*$ is unknown