MATH578 - Numerical Analysis 1

Based on lectures from Fall 2025 by Prof. J.C. Nave.
Notes by Louis Meunier

Contents

§1 Polynomial Interpolation

In general, the goal of interpolation is, given a function $f(x)$ on $[a,b]$ and a series of distinct ordered points (often called *nodes* or *collocation points*) $\{x_j\}_{j=1}^n \subseteq [a,b]$, to find a polynomial $P(x)$ such that $f(x_j) = P(x_j)$ for each j.

↪**Theorem 1.1** (Existence and Uniqueness of Lagrange Polynomial): Let $f \in C[a,b]$ and $\{x_j\}$ a set of n distinct points. Then, there exists a unique $P(x) \in \mathbb{P}_{n-1}$, the space of $n-1$-degree polynomials, such that $P(x_j) = f(x_j)$ for each j.

We call such a P the *Lagrange polynomial* associated to the points $\{x_j\}$ for f.

Proof. We define the following $n-1$ degree "fundamental polynomials" associated to $\{x_j\}$,

$$\ell_j(x) \equiv \prod_{\substack{1 \leq i \leq n \\ i \neq j}} \frac{x - x_i}{x_j - x_i}, \qquad j = 1, \ldots, n.$$

Then, one readily verifes $\ell_j(x_i) = \delta_{ij}$, and that the distinctness of the nodes guarantees the denominator in each term of the product is nonzero. Define

$$P(x) = \sum_{j=1}^{n} f(x_j)\ell(x),$$

which, being a linear combination of $n-1$ degree polynomials is also in $\mathbb{P}_{n-1}$. Moreover,

$$P(x_i) = \sum f(x_j)\delta_{i,j} = f(x_i),$$

as desired.

For uniqueness, suppose $\overline{P}$ another $n-1$ degree polynomial satisfying the conditions of the theorem. Then, $q(x) \equiv P(x) - \overline{P}(x)$ is also a degree $n-1$ polynomial with $q(x_i) = 0$ for each $i = 1, \ldots n$. Hence, q a polynomial with more distinct roots than its degree, and thus it must be identically zero, hence $P = \overline{P}$, proving uniqueness. ■

↪**Theorem 1.2** (Interpolation Error): Suppose $f \in C^n[a,b]$, and let $P(x)$ be the Lagrange polynomial for a set of n points $\{x_j\}$, with $x_1 = a, x_n = b$. Then, for each $x \in [a,b]$, there is a $\xi \in [a,b]$ such that

$$f(x) - P(x) = \frac{f^{(n)}(\xi)}{n!}(x - x_1)\cdots(x - x_n).$$

Moreover, if we put $h := \max_i (x_{i+1} - x_i)$, then

$$\|f - P\|_\infty \leq \frac{h^n}{4n}\|f^{(n)}\|_\infty.$$

Proof. We prove the first identity, and leave the second "Moreover" as a homework problem. Notice that it holds trivially for $x = x_j$ for any j, so assume $x \neq x_j$ for any j, and define the function

$$g(t) := f(t) - P(t) - \omega(t)\frac{f(x) - P(x)}{\omega(x)}, \qquad \omega(t) := (t - x_1)...(t - x_n) \in \mathbb{P}_n[t].$$

Then, we observe the following:

- $g \in C^n[a, b]$
- $g(x) = 0$
- $g\left(t = x_j\right) = 0$ for each j

Recall that by Rolle's Theorem, if a C^1 function has $\geq m$ roots, then its derivative has $\geq m - 1$ roots. Thus, applying this principle inductively to $g(t)$, we conclude that $g^{(n)}(t)$ has at least one root. Take ξ to be such a root. Then, one readily verifies that $P^{(n)} \equiv 0$ and $\omega^{(n)} \equiv n!$ (using polynomial properties), from which we may use the fact that $g^{(n)}(\xi) = 0$ to simplify to the required identity. ■

Remark 1.1: In general, larger n leads to smaller maximum step size h. However, it is *not* true that $n \to \infty$ implies $P \to f$ in L^∞. From the previous theorem, one would need to guarantee $\left\|f^{(n)}\right\| \to 0$ (or at least, doesn't grow faster than $\frac{h^n}{4n}$), which certainly won't hold in general; we have no control on the nth-derivative of an arbitrarily given function. However, we can try to optimize our choice of points $\left\{x_j\right\}$ for a given j.

We switch notation for convention's sake to $n + 1$ points x_j. Our goal is the optimization problem

$$\min_{x_j} \max_{x \in [a,b]} \left|\prod_j \left(x - x_j\right)\right|,$$

the only term in the error bound above that we have control over. Remark that we can expand the product term:

$$\prod_j \left(x - x_j\right) = x^{n+1} - r(x),$$

where $r(x) \in \mathbb{P}_n$. So, really, we equivalently want to solve the problem

$$\min_{r \in \mathbb{P}_n} \left\|x^{n+1} - r(x)\right\|_\infty,$$

namely, what n-degree polynomial minimizes the max difference between x^{n+1}?

↪**Theorem 1.3** (De la Vallé-Poussin Oscillation Theorem): Let $f \in C([a, b])$, and suppose $r \in \mathbb{P}_n$ for which there exists $n+2$ distinct points $\{x_j\}$ such that $a \leq x_0 < \cdots < x_{n+1} \leq b$ at which the error $f(x) - r(x)$ "oscillate" sign, i.e.

$$\operatorname{sign}\left(f\left(x_j\right) - r\left(x_j\right)\right) = -\operatorname{sign}\left(f\left(x_{j+1}\right) - r\left(x_{j+1}\right)\right).$$

Then,

$$\min_{P \in \mathbb{P}_n} \|f - P\|_\infty \geq \min_{0 \leq j \leq n+1} |f\left(x_j\right) - r\left(x_j\right)|.$$

↪**Definition 1.1** (Chebyshev Polynomial): The *degree n Chebyshev polynomial*, defined on $[-1, 1]$, is defined by

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

Remark 1.2: The fact that T_n actually is a polynomial follows from the double angle formula for cos, which says

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

In the context of T_n, this implies that for any $n \geq 1$, the recursive formula

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

This formula with a simple induction argument proves that each T_n a polynomial, with for instance $T_0(x) = 1, T_1(x) = x, T_2(x) = 2x^2 - 1$ and so on.

↪**Proposition 1.1**: $\{T_n\}$ are orthogonal with respect to the inner product given by

$$(f, g) := \int_{-1}^{1} f(x)g(x)\omega_2(x)\, \mathrm{d}x,$$

where $\omega_2(x) := \left(1 - x^2\right)^{1/2}$.

Remark 1.3: Defining similar *weight* functions by $\omega_n(x) := (1 - x^n)^{1/n}$, one can derive a more general class of polynomials called *Geigenbauer polynomials*, which are respectively orthogonal with respect to $\int \cdot \cdot \omega_n$.

↪**Proposition 1.2** (Some Properties of T_n):
- $|T_n(x)| \leq 1$ on $[-1,1]$
- The roots of $T_n(x)$ are the n points

$$\xi_j := \cos\left(\frac{(2j-1)\pi}{2n}\right), \qquad j = 1, ..., n.$$

- For $n \geq 1$, $|T_n(x)|$ is maximal on $[-1,1]$ at the $n+1$ points

$$\eta_j := \cos\left(\frac{j\pi}{n}\right), \qquad j = 0, ..., n,$$

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

Note too that $T_{n+1}(x)$ has leading coefficient 2^n, which can be seen by the recursive formula above; define the *normalized* Chebyshev polynomials by $\hat{T}_{n+1}(x) := 2^{-n}T_{n+1}(x)$. Thus, we may write

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

with $r_n(x) \in \mathbb{P}_n$. It follows for one that

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

Moreover, we know that at the $n+2$ points η_j, we have

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

Namely, because of the inclusion of $(-1)^j$ term, this means that $\hat{T}_{n+1}(x)$ oscillates sign between the η_j points, which fulfils the condition stated in the Oscillation Theorem. Thus, these observations readily imply the following result, settling our original question on optimizing locations of interpolation points for Lagrange interpolation:

↪**Theorem 1.4** (Optimal Approximation of x^{n+1} in $\mathbb{P}_n$): The optimal approximation of x^{n+1} in $\mathbb{P}_n$ on $[-1,1]$ with respect to the L^∞ norm is given by

$$r_n(x) := x^{n+1} - 2^{-n}T_{n+1}(x).$$

Thus, the optimal Lagrange interpolation points are the $n+1$ roots of $x^{n+1} - r_n(x)$, namely $\xi_j = \cos\left(\frac{(2j+1)\pi}{2n+2}\right)$ for $j = 0, ..., n$.

Remark 1.4: This, and previous results, were stated over $[-1,1]$. A linear change of coordinates transforming any closed interval to $[-1,1]$ readily leads to analgous results.

§2 Fourier Transform

Recall that the Fourier transform of a (Lebesgue) measurable function $u(x)$ on $\mathbb{R}$ is defined

$$(\mathcal{F}u)(\xi) = \hat{u}(\xi) = \int_{\mathbb{R}} e^{-i\xi x} u(x) \, \mathrm{d}x.$$

↪**Theorem 2.1**: Let $u \in L^2(\mathbb{R})$. Then,

1. $\hat{u} \in L^2$
2. the *inversion* formula holds, ie $u(x) = \int_{\mathbb{R}} \hat{u}(\xi) e^{i\xi x} \, dx = \left(\mathcal{F}^{\{-1\}} u\right)(x)$
3. $\|\hat{u}\|_2 = \sqrt{2\pi} \|u\|_2$
4. for $u \in L^2, v \in L^1, u * v \in L^2$, and $\widehat{u * v} = \hat{u}\hat{v}$.

↪**Theorem 2.2** (Further Properties of Fourier Transform): Let $u, v \in L^2$. Then,

1. $\mathcal{F}$ is linear over $\mathbb{R}$
2. $\mathcal{F}(u(\cdot + x_0))(\xi) = e^{i\xi x_0} \hat{u}(x_0)$
3. $\mathcal{F}\left(e^{i\xi_0 x} u(x)\right)(\xi) = \hat{u}(\xi - x_0)$
4. If $c \neq 0$, $\mathcal{F}(u(c \cdot))(\xi) = \frac{\hat{u}\left(\frac{\xi}{c}\right)}{c}$
5. $\mathcal{F}(\overline{u})(\xi) = \overline{\hat{u}(-\xi)}$
6. if u_x exists and is in L^2, then

$$\mathcal{F}(u_x)(\xi) = i\xi\hat{u}(\xi).$$

   By extension, if $\partial_\alpha u \in L^2$, then $\widehat{\partial_\alpha u}(\xi) = (i\xi)^\alpha \hat{u}(\xi)$
7. $\left(\mathcal{F}^{-1} u\right)(\xi) = \frac{1}{2\pi} \hat{u}(-\xi)$.

In a sense, 6. implies a duality between the smoothness of $u(x)$ and rapid decay (as $|\xi| \to \infty$) of $\hat{u}(x)$; 7. indicates that the same analogy holds switching the roles of u and $\hat{u}$. We make this more precise.

↪**Definition 2.1** (Bounded Variation): We say a function u on $\mathbb{R}$ is of *bounded variation* or write $\in$ BV if there exists a constant M such that for any finite integer m and collection of points $x_0 < x_1 < \ldots < x_m$,

$$\sum_{j=1}^{m} \left| u\left(x_j\right) - u\left(x_{j-1}\right) \right| \leq M.$$

In a sense, this notion of BV captures a notion of "limited oscillation".

↪**Theorem 2.3**: Let $u \in L^2$. Then:

1. If u has $p - 1$ continuous derivatives in L^2 and its pth derivative is in BV, then

$$\hat{u}(\xi) = O\left(|\xi|^{-p-1}\right).$$

2. If u has infinitely many derivatives all in L^2, then

$$\hat{u}(\xi) = O\left(|\xi|^{-M}\right), \qquad \forall M \geq 1.$$

§2.1 Discrete Fourier Transform

Let $h > 0$ be a *step size*. Let $x_j = jh$ for $j \in \mathbb{Z}$. We write $v = \left\{v_j\right\}_{j \in \mathbb{Z}}$ for discrete approximations of a function u on the grid $\left\{x_j\right\}_{j \in \mathbb{Z}}$, i.e. $v_j \approx u\left(x_j\right)$.

The ℓ_h^2 norm is defined for such v by

$$\|v\|_2 := \left[h \sum_{j \in \mathbb{Z}} |v_j|^2 \right]^{1/2}.$$

Then, ℓ_h^2 is defined as the space of such sequences v such that this norm is finite. analogous definitions hold for other ℓ_h^p spaces and norms.

↪**Proposition 2.1** (Nesting): $\ell_h^p \subset \ell_h^q$ for each $q \geq p$.

Remark 2.1: Note that the analogous result to this does *not* hold for L^p spaces (unless restricted to a compact domain).

We define the convolution of two sequences v, w by the new sequence $v * w$ with entries

$$(v * w)_m = h \sum_{j \in \mathbb{Z}} v_j w_{m-j} = h \sum_{j \in \mathbb{Z}} v_{m-j} w_j.$$

For any $v \in \ell_h^2$, we define too the *semi-discrete Fourier transform* of v by

$$\hat{v}(\xi) = (\mathcal{F}_h v)(\xi) = h \sum_{j \in \mathbb{Z}} e^{-i\xi x_j} v_j, \qquad \xi \in \left[-\frac{\pi}{h}, \frac{\pi}{h} \right],$$

where we remark that $\hat{v}(\xi)$ $\frac{2\pi}{h}$-periodic (hence the domain restriction) and continuous.