Chapter 2
Approximation and Interpolation

Songting Luo

Department of Mathematics
Iowa State University

MATH 561 Numerical Analysis

Approximation of Functions

Approximation of Functions

Given a function f and a class Φ of "approximating functions" ϕ and a norm $\|\cdot\|$. A function $\hat{\phi} \in \Phi$ is called the best approximation of f from the class Φ relative to the norm $\|\cdot\|$ if

$$\|f - \hat{\phi}\| \leqslant \|f - \phi\| \text{ for all } \phi \in \Phi.$$

Depending on the linear space Φ and norm $\|\cdot\|$.

Existence?

Uniqueness?

Approximation Error?

Approximation of Functions; Examples

Depending on the linear space Φ and norm $\|\cdot\|$.
For example:

- Least Squares Approximation: the least squares problem:

$$\min_{\phi \in \Phi_n} \|\phi - f\|_{2,d\lambda}.$$

  - Existence and Uniqueness by Normal equations; Least Squares Error.

- Polynomial Interpolation: given $\{x_i\}_{i=0}^{n}$ and $\{f_i = f(x_i)\}_{i=1}^{n}$ of function f, find a polynomial $p \in \mathbf{P}_n$ s.t.,

$$p(x_i) = f_i, \; i = 0, 1, \ldots, n.$$

- Fourier series with trigonometric functions, use truncated sum to approximate the function (recall Calculus).

Approximation of Functions by Polynomials

- Polynomial Interpolation
  - Polynomial interpolation: Vandermonde method, Lagrange formula, Barycentric formula, Newton's formula; Interpolation error; Chebyshev Nodes.
  - Hermite interpolation: Newton's formula.
  - Spline interpolation/approximation, piecewise Lagrange interpolation.
- Polynomial Approximation
  - Weierstrass's Approximation Theorem (A Proof with Bernstein polynomial)
  - Characterization of best approximation; Alternant set; Remez method.
  - Least Squares Approximation; Normal equations.
  - Orthogonal polynomials: Chebyshev polynomials, Legendre polynomials.

Approximation and Interpolation by Spline Functions

Spline Functions

Given a partition (subdivision) Δ of $[a,\ b]$,

$$\Delta : a = x_1 < x_2 < \cdots < x_{n-1} < x_n = b$$

with $|\Delta| \equiv \max_{1 \leqslant i \leqslant n-1} \Delta x_i,\ \ \Delta x_i = x_{i+1} - x_i$.
Recall the spline functions of degree m and smoothness class k relative to the subdivision Δ,

$$\mathbf{S}_m^k(\Delta) = \{s : s \in C^k[a,\ b], s|_{[x_i,\ x_{i+1}]} \in \mathbf{P}_m, i = 1, 2, \ldots, n-1\}.$$

I.e., any function in $\mathbf{S}_m^k$ is piecewise polynomial of degree $\leqslant m$, and upto kth derivative is continuous everywhere including points $x_1, \ldots, x_{n-1}$ of Δ.

Approximation and Interpolation by Spline Functions

Some examples:

- $\mathbf{S}_m^{-1}$: piecewise polynomial of degree $\leqslant m$, no assumption of continuity at $x_1, \ldots, x_{n-1}$ is assumed.
- $\mathbf{S}_m^m = \mathbf{P}_m$.
- $k < m$: e.g., simplest case $m = 1,\ k = 0$, i.e., piecewise linear interpolation.

Piecewise Linear Interpolation

Interpolation by Piecewise Linear Functions

Find $s \in \mathbf{S}_1^0(\Delta)$ such that for a given function f defined on $[a,\ b]$,

$$s(x_i) = f_i \text{ where } f_i = f(x_i),\ i = 1, 2, \ldots, n.$$

The solution is given by $s(\cdot) = s_1(f;\cdot)$, on $[x_i, x_{i+1}]$:

$$s_1(f;x) = f_i + (x - x_i)[x_i, x_{i+1}]f \text{ for } x_i \leqslant x \leqslant x_{i+1},\ \ i = 1, 2, \ldots, n-1.$$

I.e., on each subinterval $[x_i, x_{i+1}]$, s is a linear function. The interpolation error is (from previous results with Newton's form)

$$f(x) - s_1(f;x) = (x - x_i)(x - x_{i+1})[x_i, x_{i+1}, x]f \text{ for } x \in [x_i, x_{i+1}]$$

If $f \in C^2[a,\ b]$,

$$|f(x) - s_1(f;x)| \leqslant \frac{(\Delta x_i)^2}{8} \max_{[x_i, x_{i+1}]} |f''|,\ \ x \in [x_i, x_{i+1}].$$

Piecewise Linear Functions; Interpolation Error

Interpolation error:

$$\|f(\cdot) - s_1(f;\cdot)\|_\infty \leqslant \frac{1}{8}|\Delta|^2\|f''\|_\infty$$

Furthermore, the piecewise linear interpolation is nearly optimal:

$$dist_\infty(f, \mathbf{S}_1^0) \leqslant \|f(\cdot) - s_1(f;\cdot)\|_\infty \leqslant 2dist_\infty(f, \mathbf{S}_1^0)$$

where

$$dist_\infty(f, \mathbf{S}) \equiv \inf_{s\in\mathbf{S}} \|f(\cdot) - s\|_\infty$$

is the best approximation to f from $\mathbf{S}$.

Basis for $\mathbf{S}_1^0(\Delta)$

Dimension of $\mathbf{S}_1^0(\Delta)$: n.

A basis: for $i = 1, \dots, n$, (denote $x_0 = x_1$ and $x_{n+1} = x_n$)

$$B_i(x) = \begin{cases} \dfrac{x - x_{i-1}}{x_i - x_{i-1}} \text{ if } x_{i-1} \leqslant x \leqslant x_i, \\ \dfrac{x_{i+1} - x}{x_{i+1} - x_i} \text{ if } x_i \leqslant x \leqslant x_{i+1}, \\ 0 \text{ otherwise.} \end{cases}$$

Clearly,

$$B_i(x_j) = \delta_{ij} = \begin{cases} 1, \text{ if } i = j, \\ 0, \text{ if } i \neq j. \end{cases}$$

And for any $s \in \mathbf{S}_1^0(\Delta)$,

$$s(x) = \sum_{i=1}^{n} s(x_i) B_i(x).$$

Least Squares Approximation over $\mathbf{S}_1^0(\Delta)$

Least Squares Approximation

Given $f \in C[a,\ b]$, find $\hat{s}_1(f;\cdot) \in \mathbf{S}_1^0(\Delta)$ such that

$$\|f - \hat{s}_1\|_2 = \min_{s \in \mathbf{S}_1^0(\Delta)} \|f - s\|_2.$$

The unique solution by Normal equations

$$\mathbf{Ac} = \mathbf{b}$$

with $\mathbf{A} = [a_{ij}] = [(B_i, B_j)]$, $\mathbf{b} = [b_i] = [(f, B_i)]$, denoted as $\hat{\mathbf{c}} = \mathbf{A}^{-1}\mathbf{b}$ or $\hat{s}_1(f;x) = \sum_{i=1}^{n} \hat{c}_i B_i(x)$. Clearly, $(B_i, B_j) = 0$ if $|i-j| > 1$, so $\mathbf{A}$ is tridiagonal, i.e.,

$$\frac{1}{6}\Delta x_{i-1}\hat{c}_{i-1} + \frac{1}{3}(\Delta x_{i-1} + \Delta x_i)\hat{c}_i + \frac{1}{6}\Delta x_i \hat{c}_{i+1} = b_i,\ \ i = 1, 2, \ldots, n.$$

The least squares approximation is nearly optimal:

$$dist_\infty(f, \mathbf{S}_1^0) \leqslant \|f(\cdot) - \hat{s}_1(f;\cdot)\|_\infty \leqslant 4 dist_\infty(f, \mathbf{S}_1^0).$$

Interpolation by Cubic Splines

Cubic Splines $\mathbf{S}_3^1(\Delta)$

Given nodes $x_1, \ldots, x_n$, and numbers $m_1, \ldots, m_n$, find $s_3(f;\cdot) \in \mathbf{S}_3^1(\Delta)$ with

$$s_3(f;\cdot)|_{[x_i,x_{i+1}]} \equiv p_i(x), \ i = 1, 2, \ldots, n-1,$$

such that $s_3'(f;x_i) = m_i, \ i = 1, \ldots, n.$

We selecting each piece p_i to be the solution of a Hermite interpolation problem: for $i = 1, 2, \ldots, n-1$,

$$p_i(x_i) = f_i, \quad p_i(x_{i+1}) = f_{i+1},$$

$$p_i'(x_i) = m_i, \quad p_i'(x_{i+1}) = m_{i+1}.$$

The cubic splines depend on the choices of $m_1, \ldots, m_n$. Different approaches used to determined $m_1, \ldots, m_n$ result in different cubic splines (discussed later).

Newton's Formula; In general

Cubic Splines $\mathbf{S}_3^1(\Delta)$

Each piece p_i is given by

- in Newton's form

$$p_i(x) = f_i + (x - x_i)m_i + (x - x_i)^2 \frac{[x_i, x_{i+1}]f - m_i}{\Delta x_i} + (x - x_i)^2 (x - x_{i+1}) \frac{m_{i+1} + m_i - 2[x_i, x_{i+1}]f}{(\Delta x_i)^2}$$

- in Taylor's form

$p_i(x) = c_{i,0} + c_{i,1}(x - x_i) + c_{i,2}(x - x_i)^2 + c_{i,3}(x - x_i)^3$, with

$$c_{i,0} = f_i; \quad c_{i,1} = m_i; \quad c_{i,2} = \frac{[x_i, x_{i+1}]f - m_i}{\Delta x_i} - c_{i,3}\Delta x_i;$$

$$c_{i,3} = \frac{m_{i+1} + m_i - 2[x_i, x_{i_1}]f}{(\Delta x_i)^2}$$

Possible Choices of $\{m_i\}$

- Piecewise cubic Hermite interpolation: for $i = 1, \ldots, n$,

$$m_i = f'(x_i).$$

- Cubic spline interpolation: $s_3(f;\cdot) \in \mathbf{S}_3^2(\Delta)$, i.e., enforcing,

$$p''_{i-1}(x_i) = p''_i(x_i),\ i = 2, 3, \ldots, n-1.$$

then from Taylor's form, we have

$$2c_{i-1,2} + 6_{i-1,3} \cdot \Delta x_{i-1} = 2c_{i,2},\ i = 2, 3, \ldots, n-1,$$

which can reformulated as a linear system for $m_1, \ldots, m_n$:

$$(\Delta x_i)m_{i-1}+2(\Delta x_{i-1}+\Delta_i)m_i+(\Delta x_{i-1})m_{i+1} = b_i,\ i = 2, 3, \ldots, n-1,$$

with $b_i = 3\{(\Delta x_i)[x_{i-1}, x_i]f + (\Delta x_{i-1})[x_i, x_{i+1}]f\}$.
?Only $n-2$ equations for n unknowns $m_1, \ldots, m_n$? Not Enough!
Need m_1, m_n in some way!

Cubic Spline Interpolation

- Cubic spline interpolation: $s_3(f;\cdot) \in \mathbf{S}_3^2(\Delta)$, i.e., enforcing,

$$p''_{i-1}(x_i) = p''_i(x_i),\ i = 2, 3, \ldots, n-1.$$

Once m_1, m_n are chosen, the linear system on $m_1, \ldots, m_n$ can be solved easily by Gauss Elimination.

  - Complete (clamped) splines:

$$m_1 = f'(a), m_n = f'(b)$$

  - Matching of the second derivative at the endpoints:

$$s''_3(f;a) = f''(a), s''_3(f;b) = f''(b)$$

  - Natural cubic splines:

$$s''(f;a) = s''(f;b) = 0$$

  - Not-a-knot spline.

Minimality Properties of Cubic Splines

Complete and natural splines have interesting optimality properties.
Subdivision Δ:

$$\Delta : a = x_1 < x_2 < \cdots < x_{n-1} < x_n = b$$

Subdivision Δ':

$$\Delta' : a = x_0 = x_1 < x_2 < \cdots < x_{n-1} < x_n = x_{n+1} = b$$

Minimality Properties of Cubic Splines

Theorem (Complete Cubic Spline Interpolant)

For any function $g \in C^2[a,\ b]$ *that interpolates* f *on* Δ'*, there holds*

$$\int_a^b [g''(x)]^2 dx \geqslant \int_a^b [s''_{compl}(f;x)]^2 dx$$

with equality iff $g(\cdot) = s_{compl}(f;\cdot)$.

Theorem (Natural Cubic Spline Interpolant)

For any function $g \in C^2[a,\ b]$ *that interpolates* f *on* Δ *(not* Δ'*), there holds*

$$\int_a^b [g''(x)]^2 dx \geqslant \int_a^b [s''_{nat}(f;x)]^2 dx$$

with equality iff $g(\cdot) = s_{nat}(f;\cdot)$.

Note: $\int_a^b [s''_{compl}(f;x)]^2 dx \geqslant \int_a^b [s''_{nat}(f;x)]^2 dx$.

Weierstrass's Theorem; Bernstein Polynomials

Theorem (Weierstrass's Approximation Theorem)

If $f(x) \in C[a,b]$, then given $\epsilon > 0$, we can find $p(x)$ such that

$$\sup |f(x) - p(x)| < \epsilon.$$

An alternative statement of it is that a continuous function is the sum of a uniformly convergent series of polynomials. For let $p_{n_1}(x), p_{n_2}(x), \cdots (n_1 \leqslant n_2 \leqslant \cdots)$ be polynomials corresponding to $\epsilon, \epsilon/2, \ldots, \epsilon/2^n, \ldots,$ Then the series

$$p_{n_1}(x) + \{p_{n_2}(x) - p_{n_1}(x)\} + \cdots$$

converges uniformly to $f(x)$.

Proof by Bernstein polynomials.

Bernstein Polynomial

Definition

Write $l_{n,m}(x) = \left(\begin{array}{c} n \\ m \end{array} \right) x^m (1-x)^{n-m},\ 0 \leqslant m \leqslant n$. The nth Bernstein polynomials of $f(x)$ in $(0,\ 1)$ is defined to be

$$B_n(x) = B_n(f;x) = \sum_{m=0}^{n} f(m/n) l_{n,m}(x).$$

$B_n(x)$ has degree n (at most).

Theorem

Let $f \in C[0,1]$, then $B_n(x) \to f(x)$ uniformly as $n \to \infty$.

The uniform convergence can be extended to any interval $[a, b]$.

Bernstein Polynomial; Proof

Lemma

Denote $f_0(x) = 1,\ f_1(x) = x,\ f_2(x) = x^2$.

- $B_n(f_0) = f_0$, $B_n(f_1) = f_1$.
- $B_n(f_2) = (1 - \frac{1}{n})f_2 + \frac{1}{n}f_1$, hence $B_n(f_2) \to f_2$ uniformly as $n \to \infty$.
- $\sum_{k=0}^{n}(\frac{k}{n} - x)^2 \begin{pmatrix} n \\ k \end{pmatrix} x^k(1-x)^{n-k} = \frac{x(1-x)}{n} \leqslant \frac{1}{4n}$, if $0 \leqslant x \leqslant 1$.
- Given $\delta > 0$ and $0 \leqslant x \leqslant 1$, let F denote the set of k in $\{0, \ldots, n\}$ for which $|k/n - x| \geqslant \delta$. Then $\sum_{k \in F} \begin{pmatrix} n \\ k \end{pmatrix} x^k(1-x)^{n-k} \leqslant \frac{1}{4n\delta^2}$.

Proof of Weierstrass's Theorem

Let $f \in C[0,1]$, and $\delta > 0$. There is a $\delta > 0$ such that $|f(x) - f(y)| < \epsilon/2$ whenever $|x - y| < \delta$. We know $l_{n,k} \geqslant 0$ and $\sum_{k=0}^{n} l_{n,k} = 1$. Then,

$$
\begin{aligned}
|f(x) - B_n(f)(x)| &= |f(x) - \sum_{k=0}^{n} \left(\begin{array}{c} n \\ k \end{array} \right) f(k/n) x^k (1-x)^{n-k}| \\
&= | \sum_{k=0}^{n} \left(\begin{array}{c} n \\ k \end{array} \right) (f(x) - f(k/n)) x^k (1-x)^{n-k}| \\
&\leqslant \sum_{k=0}^{n} |f(x) - f(k/n)| \left(\begin{array}{c} n \\ k \end{array} \right) x^k (1-x)^{n-k}.
\end{aligned}
$$

Now fix n (to be specified later) and let F denote the set of k in $\{0, \ldots, n\}$ for which $|(k/n) - x| \geqslant \delta$. Then $|f(x) - f(k/n)| < \epsilon/2$ for $k \notin F$, while $|f(x) - f(k/n)| \leqslant 2\|f\|f$ for $k \in F$.

Proof of Weierstrass's Theorem

Thus,

$$
\begin{aligned}
&|f(x) - B_n(f)(x)| \\
&\leqslant \frac{\epsilon}{2} \sum_{k \notin F} \binom{n}{k} x^k (1-x)^{n-k} + 2\|f\| \sum_{k \in F} \binom{n}{k} x^k (1-x)^{n-k} \\
&< \frac{\epsilon}{2} \cdot 1 + 2\|f\| \cdot \frac{1}{4n\delta^2} \\
&< \epsilon, \text{ provided that } n > \|f\|/\epsilon\delta^2.
\end{aligned}
$$

Characterization of Best Approximation

We denote for any $f \in C[a, b]$,

$$E_n(f) = \inf_{p \in \mathbf{P}_n} \|f - p\|, \forall n \geqslant 0.$$

Clearly,

$$E_0(f) \geqslant E_1(f) \geqslant \cdots \geqslant E_n(f) \geqslant \cdots ,$$

by Weierstrass's Theorem

$$\lim_{n \to \infty} E_n(f) = 0, \forall f \in C[a, b].$$

Definition (Best Uniform Approximation)

A best uniform approximation of a given $f \in C[a, b]$ in $\mathbf{P}_n$ is a polynomial $p_n \in \mathbf{P}_n$ that satisfies $\|f - p_n\| = \min_{p \in \mathbf{P}_n} \|f - p\|$.

A best uniform approximation is also called a minimax approximation, because $\max_{a \leqslant x \leqslant b} |f(x) - p_n(x)| = \min_{p \in \mathbf{P}_n} \max_{a \leqslant x \leqslant b} |f(x) - p(x)|$.

Existence of Best Uniform Approximation

Theorem (Existence)

For any $f \in C[a,b]$ and any $n \geqslant 0$, there exits a best uniform approximation of f in $\mathbf{P}_n$.

Proof. Let $f \in C[a,b]$ and $n \geqslant 0$. For any $\mathbf{c} = (c_0, \ldots, c_n) \in \mathbf{R}^{n+1}$, define a $p_c \in \mathbf{P}_n$ as $p_c(x) = \sum_{k=0}^{n} c_k x^k$. Define $F : \mathbf{R}^{n+1} \to \mathbf{R}$ by

$$F(\mathbf{c}) = \|f - p_c\| = \max_{a \leqslant x \leqslant b} |f(x) - \sum_{k=0}^{n} c_k x^k\|.$$

The assertion of the theorem is equivalent to the existence of $\mathbf{c} \in \mathbf{R}^{n+1}$ s.t.

$$F(\mathbf{c}) = \min_{\mathbf{d} \in \mathbf{R}^{n+1}} F(\mathbf{d}).$$

Let $m = \inf_{\mathbf{d} \in \mathbf{R}^{n+1}} F(\mathbf{d})$. Since $\lim_{\mathbf{d} \to \infty} F(\mathbf{d}) = \infty$, there exits $R > 0$ s.t. $F(\mathbf{d}) > m$ if $\|\mathbf{d}\| > R$. Hence $m = \inf_{\|\mathbf{d}\| \leqslant R} F(\mathbf{d})$. By continuity of F, this minimum is obtained on $\{\mathbf{d} \in \mathbf{R}^{n+1} : \|\mathbf{d}\| \leqslant R\}$.

Best Uniform Approximation; Alternating Set

Theorem (The Chebyshev Alternation Theorem)

Let $f \in C[a,b]$ *and* $f \notin \mathbf{P}_n$. *Then* $p \in \mathbf{P}_n$ *is a best uniform approximation if and only if* $f - p$ *achieves its maximum magnitude at* $n+2$ *points with alternating signs, i.e., there exist* $n+2$ *points* $\{x_1 < x_2 < \cdots < x_{n+2}\}$ *in* $[a,b]$ *such that*

$$|f(x_k) - p(x_k)| = \|f - p\|, \; k = 1, \ldots, n+2,$$

$$(f(x_k) - p(x_k))(f(x_{k+1}) - p(x_{k+1})) < 0, \quad k = 1, \ldots, n+1.$$

Definition (Change of Sign)

A function $g : (a,b) \to \mathbf{R}$ changes its sign at a point z, if there exists $\epsilon > 0$ with $(z-\epsilon, z+\epsilon) \subset (a,b)$ s.t.

- $g(x) \geqslant (\leqslant) 0$ for $x \in (z-\epsilon, z)$, and $g(x) \leqslant (\geqslant) 0$ for $x \in (z, z+\epsilon)$.
- both one-side limits $g(z-)$ and $g(z+)$ exist and they are not equal.

Proof of the Chebyshev Alternation Theorem

It is equivalent to proving:

Let $f \in C[a, b]$ but $f \notin \mathbf{P}_n$. Then the zero polynomial $0 \in \mathbf{P}_n$ is a best uniform approximation of f in $\mathbf{P}_n$ if and only if f achieves its maximum magnitude at $n+2$ points $\{x_1, \dots, x_{n+2}\}$ in $[a, b]$ with alternating signs.

Proof by contradiction.

Uniqueness of Best Uniform Approximation

Theorem (Uniqueness)

For any $f \in C[a,b]$ and $n \geqslant 0$, the best approximation of f is unique.

Proof by contradiction. Suppose p, q are both best approximations. Then $r = (p+q)/2$ is a best approximation. By Chebyshev Alternation Theorem, $|f - r|$ attains maximum at $\{x_1, \ldots, x_{n+2}\}$.
Assume $f(x_k) - r(x_k) = E_n(f)$ for some k, which implies $f(x_k) - (p(x_k) + q(x_k))/2 = E_n(f) = \|f - p\| \geqslant f(x_k) - p(x_k)$. Thus $p(x_k) \leqslant q(x_k)$. Similarly, $q(x_k) \geqslant p(x_k)$. So $p(x_k) = q(x_k)$.
Similarly, assume $f(x_j) - r(x_j) = -E_n(f)$ for some j, we can prove $p(x_j) = q(x_j)$.
To summarize, $p = q$ at $n+2$ points $\{x_1, \ldots, x_{n+2}\}$, hence $p = q$ in $\mathbf{P}_n$.

Chebyshev Polynomials

Chebyshev polynomials of first kind.

- $T_n(x) = \cos(n\cos^{-1}x),\ x \in [-1,1]$. $(\cos n\theta = T_n(\cos\theta))$
- $T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x)$. $T_0(x) = 1,\ T_1(x) = x$. $(\cos(n+1)\theta + \cos(n-1)\theta = 2\cos\theta\cos n\theta)$
- Zeros of T_n: $x_k^{(n)} = \cos\theta_k^{(n)},\ \theta_k^{(n)} = \frac{2k-1}{2n}\pi,\ k = 1,2,\ldots,n$.
- Extrema of T_n: $y_k^{(n)} = \cos\eta_k^{(n)},\ \eta_k^{(n)} = k\frac{\pi}{n},\ k = 0,1,2,\ldots,n$. So $\mathring{T}_n$ obtains maximum magnitude at $n+1$ points with alternating signs. That is $x^n - (x^n - \mathring{T}_n)$ obtains maximum magnitude at $n+1$ points with alternating signs. By Chebyshev Alternation Theorem, $(x^n - \mathring{T}_n)$ is the best uniform approximation of x^n in $\mathbf{P}_{n-1}$. I.e.,

$$1/2^{n-1} = \max_{-1\leqslant x\leqslant 1} |\mathring{T}_n| = \max_{-1\leqslant x\leqslant 1} |x^n - (x^n - \mathring{T}_n)|$$

$$= \min_{p_{n-1}\in\mathbf{P}_{n-1}} \max_{-1\leqslant x\leqslant 1} |x^n - p_{n-1}| = \min_{\mathring{p}_n\in\mathring{\mathbf{P}}_n} \max_{-1\leqslant x\leqslant 1} |\mathring{p}_n|$$

Properties of Chebyshev Polynomial of First Kind

- Orthogonality with weight $w(x) = 1/\sqrt{1-x^2}$:

$$\int_{-1}^{1} T_k(x)T_l(x)\frac{dx}{\sqrt{1-x^2}} = \begin{cases} 0, \text{ if } k \neq l \\ \pi, \text{ if } k = l = 0 \\ \frac{\pi}{2}, \text{ if } k = l > 0 \end{cases}$$

- Best uniform approximation (by Chebyshev's Theorem); and best least squares approximation

$$\int_{-1}^{1} \frac{1}{\sqrt{1-x^2}}[\mathring{T}_n(x)]^2 dx = \min_{\mathring{p}_n \in \mathring{\mathbf{P}}_n} \int_{-1}^{1} \frac{1}{\sqrt{1-x^2}}[\mathring{p}_n(x)]^2 dx = 2^{1-2n}\pi.$$

- Differential equation:
$(1-x^2)T_n''(x) - xT_n'(x) + n^2 T_n(x) = 0, \quad n = 0, \ldots.$
- Rodrigue's formula:
$T_n(x) = \frac{(-1)^n}{(2n-1)!!}(1-x^2)^{1/2}\frac{d^n}{dx^n}(1-x^2)^{n-1/2}, \; n = 0, \ldots.$

Orthogonal Polynomials

Recall inner product w.r.t. weight $w(x)$ on $[a, b]$:

$$(f, g) = \int_a^b w(x) f(x) g(x) dx.$$

Definition (Orthogonal Polynomials)

A sequence of polynomials $Q_n,\ n = 0, 1, \ldots$, are called orthogonal polynomials in $L_w^2(a, b) \equiv \{s : (s, s) < \infty\}$, if

- Q_n is a polynomial of degree n
- $(Q_n, Q_m) = 0$ if $m \neq n$.

Lemma

- *If* $\{Q_0, \ldots, Q_n\}$ *are orthogonal, then they are linearly independent.*
- $\{Q_0, \ldots, Q_n\}$ *is a basis of* $\mathbf{P}_n$.

Properties of Orthogonal Polynomials

Theorem (Least Squares Approximation)

Let $\{Q_0, \ldots, Q_n\}$ *be orthogonal polynomials. Then the least squares approximation of a given function* f *in* $\mathbf{P}_n$ *is*

$$p_n = \sum_{k=0}^{n} \frac{(f, Q_k)}{(Q_k, Q_k)} Q_k.$$

See previous lectures for proof. Also we know,

Theorem

Three-term Recurrence for Orthogonal Polynomials Let $Q_0(x) = 1, Q_1(x) = x - a_1,\ Q_n(x) = (x - a_n)Q_{n-1}(x) - b_n Q_{n-2}(x),\ n = 2, \ldots,$ *with* $a_n = (xQ_{n-1}, Q_{n-1})/(Q_{n-1}, Q_{n-1}),\ n = 1, 2, \ldots$ *and* $b_n = (Q_{n-1}, Q_{n-1})/(Q_{n-2}, Q_{n-2}), n = 2, \ldots,$ *then* $\{Q_n,\ n = 0, \ldots\}$ *are orthogonal.*

Properties of Orthogonal Polynomials

Theorem (Minimization)

Let $\{Q_n,\ n = 0, \ldots\}$ *be orthogonal polynomials. Suppose* $n \geqslant 1$ *and* $Q_n \in \mathring{\mathbf{P}}_n$*, then* Q_n *is the unique polynomial in* $\mathring{\mathbf{P}}_n$ *s.t.,*

$$\|Q_n\| = \min_{q_n \in \mathring{\mathbf{P}}_n} \|q_n\|.$$

Proof. $Q_n(x) = x^n - q_{n-1}(x)$ for some $q_{n-1} \in \mathbf{P}_{n-1}$. Then by orthogonality,

$$0 = (Q_n, q) = (x^n - q_{n-1}(x), q), \forall q \in \mathbf{P}_{n-1}.$$

This implies q_{n-1} is the unique least-squares approximation of x^n in $\mathbf{P}_{n-1}$, which is equivalent to the assertion of the theorem.

Properties of Orthogonal Polynomials

Theorem (Minimization)

Let $\{Q_n,\ n = 0, \ldots\}$ *be orthogonal polynomials. Suppose* $n \geqslant 1$ *and* $Q_n \in \mathring{\mathbf{P}}_n$, *then* Q_n *is the unique polynomial in* $\mathring{\mathbf{P}}_n$ *s.t.,*

$$\|Q_n\| = \min_{q_n \in \mathring{\mathbf{P}}_n} \|q_n\|.$$

Proof. $Q_n(x) = x^n - q_{n-1}(x)$ for some $q_{n-1} \in \mathbf{P}_{n-1}$. Then by orthogonality,

$$0 = (Q_n, q) = (x^n - q_{n-1}(x), q), \forall q \in \mathbf{P}_{n-1}.$$

This implies q_{n-1} is the unique least-squares approximation of x^n in $\mathbf{P}_{n-1}$, which is equivalent to the assertion of the theorem.

Properties of Orthogonal Polynomials

Theorem (Uniqueness of Orthogonal Polynomials)

If $\{P_n\}$ and $\{Q_n\}$ are two systems of orthogonal polynomials in $L^2_w(a,b)$, then for each $n \geqslant 0$, there exists $c_n \in \mathbf{R}$ with $c_n \neq 0$ s.t., $P_n = c_n Q_n$.

Proof. Let α_n and β_n be the leading coefficients of P_n and Q_n, resp. Then $P_n/\alpha_n = Q_n/\beta_n$ by last theorem. DONE.

Theorem (Zeros of Orthogonal Polynomials)

Let $\{Q_n,\ n = 0, \ldots\}$ be orthogonal polynomials in $L^2_w(a,b)$. Then for $n \geqslant 1$, Q_n has exactly n simple roots in (a,b).

Proof by contradiction. We know $\int_a^b w(x)Q_n(x)dx = 0$, so Q_n changes sign in (a,b) at least once. Suppose it changes sign $k \leqslant n-1$ times at $x_1 < \cdots < x_k$. Define $p(x) = (x-x_1)\cdots(x-x_k)$, so $(Q_n, q) \neq 0$ since they change signs at same points. This is a contradiction to $(Q_n, q) = 0, \forall q \in \mathbf{P}_{n-1}$.

Legendre Polynomials

The Legendre polynomials $P_n \in \mathbf{P}_n, n = 0, \ldots$, are the unique orthogonal polynomials in $L^2(-1,1)$ that are normalized by

$$P_n(1) = 1, \forall n \geqslant 0.$$

Rogrigues' formula:

$$P_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n}[(x^2-1)^n], \; n = 0, \ldots.$$

P_n has degree n. If n is odd (even), P_n is an odd (even) polynomial. E.g.,

$$P_0(x) = 1, P_1(x) = x, P_2(x) = \frac{3}{2}x^2 - \frac{1}{2}, P_3(x) = \frac{5}{2}x^3 - \frac{3}{2}x$$

Properties of Legendre Polynomials

- Orthogonality: $\int_{-1}^{1} P_m(x)P_n(x)dx = \begin{cases} 0 & \text{if } m \neq n \\ 2/(2n+1) & \text{if } m = n. \end{cases}$
- Recurrence: $P_0(x) = 1, P_1(x) = x,$

$$(n+1)P_{n+1}(x) - (2n+1)xP_n(x) + nP_{n-1}1(x) = 0, \;\; n = 1, \ldots .$$

- For each $n \geqslant 1$, P_n has n simple roots in $(-1, 1)$.
- Least squares approximation: for $n \geqslant 1$, $\mathring{P}_n = [2^n(n!)^2/((2n)!)]P_n \in \mathbf{P}_n$ is the unique polynomial in $\mathring{\mathbf{P}}_n$ s.t.

$$\|\mathring{P}_n\|_2 = \frac{2^n(n!)^2}{(2n)!}\sqrt{2/(2n+1)} = \min_{\mathring{p}_n \in \mathring{\mathbf{P}}_n} \|\mathring{p}_n\|_2$$

- Differential equation: for $n \geqslant 0$, $(1-x^2)P_n''(x) - 2xP_n'(x) + n(n+1)P_n(x) = 0.$

Uniform Approximation by Trigonometric Polynomials

The trigonometric polynomials, i.e., functions of the form

$$P(x) = \sum_{n=-N}^{N} c_n e_n(x)$$

with $e_n(x) = exp(i2\pi nx)$ are dense in the space of periodic continuous functions.

Theorem (Approximation of Continuous Periodic Functions)

Let $f(x)$ be a complex-valued continuous function on **R** *that is 1-periodic, and let $\epsilon > 0$, then there exists a trigonometric polynomial $P(x)$ s.t., $\|f - P\|_\infty < \epsilon$.*

Proof by Weierstrass's Approximation theorem since we can pass from $[0, 1]$ to the circle using transformation $x \mapsto exp(i2\pi x)$.

Expansions in Legendre Polynomials

The polynomials $\{P_n, n = 0, \ldots\}$ form a complete set on the interval $[-1, 1]$, and any piecewise smooth function may be expanded in a series of the polynomials, i.e.,

$$f(x) = \sum_{n=0}^{\infty} c_n P_n(x), \text{ where } c_n = \frac{2n+1}{2} \int_{-1}^{1} f(x) P_n(x) dx.$$

The series will converge at each point to the usual mean of the right and left limits.

Chebyshev Polynomials

Discrete Orthogonality Relation

- Withe the zeros of $T_{n+1}(x)$ as nodes: let $n > 0, r, s \leqslant n$, and let $x_j = \cos((j + 1/2)\pi/(n + 1))$. Then

$$\sum_{j=0}^{n} T_r(x_j)T_s(x_j) = K_r\delta_{rs},$$

  where $K_0 = n + 1$ and $K_r = (n + 1)/2$ when $1 \leqslant r \leqslant n$.

- With extrema of $T_n(x)$ as nodes: let $n > 0, r, s \leqslant n$, and $x_j = \cos(\pi j/n)$, then

$$\sum_{j=0}^{n}{}'' T_r(x_j)T_s(x_j) = K_r\delta_{rs},$$

  where $K_0 = K_n = n$ and $K_r = n/2$ when $1 \leqslant r \leqslant n - 1$.

Computing Chebyshev Interpolation Polynomial

An alternative way to compute Lagrange polynomials at Chebyshev nodes: given $P_n \in \mathbf{P}_n$ be the Lagrange polynomial at $n+1$ zeros of $T_{n+1}(x)$. Since $\{T_k\}_{k=0}^n$ is a basis of $\mathbf{P}_n$, so

$$P_n(x) = \sum_{k=0}^{n}{}' c_k T_k(x),$$

where by the Discrete Orthogonality Relation, one can find

$$c_k = \frac{2}{n+1} \sum_{j=0}^{n} f(x_j) T_k(x_j), \ x_j = \cos((j+1/2)\pi/(n+1)).$$

Or

$$c_k = \frac{2}{n+1} \sum_{j=0}^{n} f(\cos\theta_j) \cos(k\theta_j), \ \theta_j = (j+1/2)\pi/(n+1),$$

which is a discrete cosine transform of $f(\cos\theta_j)$, $j = 0, \ldots, n$.

Expansions in Chebyshev Polynomials

Recall: Given f, Chebyshev interpolation (Lagrange interpolation at Chebyshev nodes) converges when the number of nodes tends to infinity. This leads to a representation of f in terms of an infinite series of Chebyshev polynomials. I.e.,

$$f(x) = \sum_{k=0}^{\infty}{}' c_k T_k(x) = \frac{1}{2} c_0 + \sum_{k=1}^{\infty} c_k T_k(x), \quad -1 \leqslant x \leqslant 1,$$

with

$$c_k = \frac{2}{\pi} \int_{-1}^{1} \frac{f(x) T_k(x)}{\sqrt{1-x^2}} dx = \frac{2}{\pi} \int_0^{\pi} f(\cos\theta) \cos(k\theta) d\theta.$$

For computing the coefficients, one needs to compute the above cosine transform.

If using truncated sum $\tau_n(x) = \sum_{k=0}^{n}{}' c_k T_k(x)$, we see

$$E_n(x) = f(x) - \tau_n(x) = \sum_{k=n+1}^{\infty} c_k T_k(x) \approx c_{n+1} T_{n+1}(x).$$

Convergence of Chebyshev Expansions

Theorem (Functions with Continuous Derivatives)

When a function f has $m+1$ continuous derivatives on $[-1,1]$, where m is a finite number, then $|f(x)-\tau_n(x)| = O(n^{-m})$ as $n \to \infty$ for all $x \in [-1,1]$.

Theorem

Analytic Functions Inside an Ellipse When a function f on $x \in [-1,1]$ can be extended to a function that is analytic inside an ellipse E_r defined by

$$E_r = \{z : |z + \sqrt{z^2-1}| = r\}, r > 1,$$

then $|f(x)-\tau_n(x)| = O(r^{-n})$ as $n \to \infty$ for all $x \in [-1,1]$.

Evaluation of a Chebyshev Sum; Clenshaw's Method

Assume $c_k, k = 0, \ldots, n$ is given, evaluate $\tau_n(x)$.

Clenshaw's Method for a Chebyshev Sum

Input: $x; c_0, c_1, \ldots, c_n$.

Output: $\tau_n(x)$.

Step 1: $b_{n+1} = 0; b_n = c_n$

Step 2: DO $r = n-1, n-2, \ldots, 1$:

$$b_r = 2xb_{r+1} - b_{r+2} + c_r.$$

Step 3: $\tau_n(x) = xb_1 - b_2 + c_0$.

Remez Method for Best Uniform Approximation

Since the best approximation is unique, we can define the operator that assigns to each continuous function f its best polynomial approximation of fixed degree p^*. This operator, although continuous, is nonlinear, and so we need iterative methods to compute p^*.

Two theorems are essential to Remez method (Evgeny Yakovlevich Remez, 1934). One is Chebyshev Alternation theorem. Another one is the following

Theorem (de La Vallée Poussin)

Let $p \in \mathbf{P}_n$ and $\{y_i\}_{i=0}^{n+1}$ be a set of $n+2$ distinct points s.t. $sign(f(y_i) - p(y_i)) = \lambda\sigma_i$ with $\sigma_i = (-1)^i$ and $\lambda = 1$ or -1 fixed. Then for any $q \in \mathbf{P}_n$, $\min_i |f(y_i) - p(y_i)| \leqslant \max_i |f(y_i) - q(y_i)|$, and in particular, $\min_i |f(y_i) - p(y_i)| \leqslant \|f - p^*\| \leqslant \|f - p\|$.

We refer to the $n+2$ points $A^* \equiv \{x_i\}_{i=0}^{n+1}$ in Chebyshev Alternation Theorem as a "reference".

Remez Method

From last theorem, we know a polynomial $p \in \mathbf{P}_n$ whose error oscillates $n+2$ times is "near-best" in the sense

$$\|f - p\| \leqslant C\|f - p^*\|, \quad C = \frac{\|f - p\|}{\min_i |f(y_i) - p(y_i)|} \geqslant 1.$$

The Remez algorithm constructs a sequence of trial references $\{A_k\}$ and trial polynomials $\{p_k\}$ that satisfy this alternation condition in such a way that $C \to 1$ as $k \to \infty$.

At the kth step the algorithm starts with a trial reference A_k and then computes a polynomial p_k s.t. $f(x_i) - p_k(x_i) = \sigma_i h_k, \ x_i \in A_k$, where $h_k = f(x_i) - p_k(x_i)$ is the levelled error. Then, a new trial reference A_{k+1} is computed from the extrema of $f - p_k$ in such a way that $|h_{k+1}| \geqslant |h_k|$ is guaranteed. This monotonic increase of the levelled error is the key observation in showing that the algorithm converges to p^*.

From a trial reference to a trial polynomial

Assume $\{\phi_j, j = 0, \ldots, n\}$ be a basis of $\mathbf{P}_n$, so

$$p(x) = \sum_{j=0}^{n} c_j \phi_j(x)$$

Then we have a linear system for $c_0, \ldots, c_n$ and h:

$$\begin{pmatrix} \phi_0(x_0) & \phi_1(x_0) & \cdots & \phi_n(x_0) \\ \phi_0(x_1) & \phi_1(x_1) & \cdots & \phi_n(x_1) \\ \vdots & \vdots & \vdots & \vdots \\ \phi_0(x_n) & \phi_1(x_n) & \cdots & \phi_n(x_n) \\ \phi_0(x_{n+1}) & \phi_1(x_{n+1}) & \cdots & \phi_n(x_{n+1}) \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} f(x_0) + \sigma_0 h \\ f(x_1) + \sigma_1 h \\ \vdots \\ f(x_n) + \sigma_n h \\ f(x_{n+1}) + \sigma_{n+1} h \end{pmatrix}$$

Choice of $\{\phi_i, \}$ is crucial.

From a trial polynomial to a new trial reference

First Remez Algorithm

Construct A_{k+1} by exchanging a point $x_{old} \in A_k$ with the global extremum x_{new} of $f - p_k$ in such a way that the alternation of signs of the error is maintained. If $x_0 < x_{new} < x_{n+1}$, then x_{old} is the closest point in A_k for which the error has the same sign as at x_{new}. If $x_{new} < x_0$ and the signs of x_{new} and x_0 coincide then x_{old} is x_0; if $x_{new} < x_0$ but the signs of x_{new} and x_0 are different, then x_{old} is x_{n+1}. Similar rules apply if $x_{new} > x_{n+1}$.

Second Remez Algorithm

Constructs the set $\tilde{A}_{k+1}$ of points in A_k and local extrema x_r of $f - p_k$ such that $|(f - p_k)(x_r)| > |h_k|$. Then, for each subset of $\tilde{A}_{k+1}$ of consecutive points with the same sign it keeps only one for which $|f - p_k|$ attains the largest value. From the resulting set, A_{k+1} is obtained by choosing $n + 2$ consecutive points that include the global extremum of $f - p_k$.