2021 Spring MAS 365
Chapter 8: Approximation Theory

Donghwan Kim

KAIST

June/1,3,8, 2021

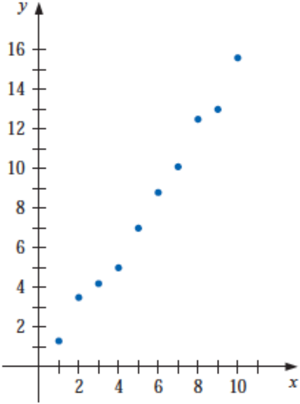
Approximation Theory

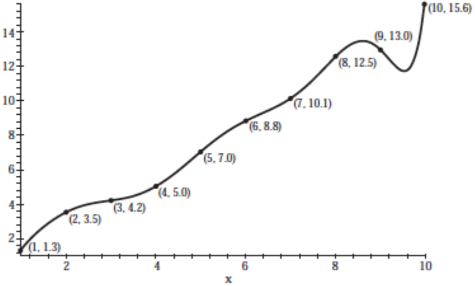
8.1 Given data, find the best function to represent the data.

8.2 Given a function, find a simpler type of function (*e.g.*, polynomial).

- We learned these in Chapter 3, and this chapter studies their limitation, and discuss other approaches.

Polynomial Approximation





- The ninth-degree interpolating polynomial is a poor predictor of information between a number of the data points.

Best Linear Approximation

- Find the "best" (in some sense) approximating line, even if it does not agree precisely with the data at any point.

1. In l_∞-norm sense, we minimize

$$E_\infty(a_0, a_1) = \max_{1 \le i \le 10} |y_i - (a_1 x_i + a_0)|,$$

   commonaly called the **minimax** problem.

2. In l_1-norm sense, we minimize

$$E_1(a_0, a_1) = \sum_{i=1}^{10} |y_i - (a_1 x_i + a_0)|,$$

   called the **absolute deviation**.

- Since the absolute-value function is not differentiable at zero, it is difficult to find the solution by setting its partial derivatives to zero.

Linear Least Squares

- In l_2-norm sense, the **least squares** approach finds the best approximating line by minimizing the sum of the squares of the differences between the y-values on the approximating line and the given y-values:

$$E_2(a_0, a_1) = \sum_{i=1}^{10} [y_i - (a_1 x_i + a_0)]^2$$

Q. What is the best among the l_1-norm, l_2-norm and l_∞ sense?

Linear Least Squares (cont'd)

- The general problem of fitting the best least squares line to a collection of data $\{(x_i, y_i)\}_{i=1}^m$ involves minimizing the total error

$$E \equiv E_2(a_0, a_1) = \sum_{i=1}^{m} [y_i - (a_1 x_i + a_0)]^2$$

with respect to a_0 and a_1.

- We thus want

$$\frac{\partial E}{\partial a_0} = 0 \quad \text{and} \quad \frac{\partial E}{\partial a_1} = 0$$

that is

$$0 = \frac{\partial}{\partial a_0} \sum_{i=1}^{m} [(y_i - (a_1 x_i - a_0)]^2 = 2 \sum_{i=1}^{m} (y_i - a_1 x_i - a_0)(-1)$$

and

$$0 = \frac{\partial}{\partial a_1} \sum_{i=1}^{m} [(y_i - (a_1 x_i - a_0)]^2 = 2 \sum_{i=1}^{m} (y_i - a_1 x_i - a_0)(-x_i).$$

Linear Least Squares (cont'd)

- Those equations simplify to the normal equations:

$$a_0 m + a_1 \sum_{i=1}^{m} x_i = \sum_{i=1}^{m} y_i$$
$$a_0 \sum_{i=1}^{m} x_i + a_1 \sum_{i=1}^{m} x_i^2 = \sum_{i=1}^{m} x_i y_i$$

- The solution to this system of equations is

$$a_0 = \frac{\sum_{i=1}^{m} x_i^2 \sum_{i=1}^{m} y_i - \sum_{i=1}^{m} x_i y_i \sum_{i=1}^{m} x_i}{m\left(\sum_{i=1}^{m} x_i^2\right) - \left(\sum_{i=1}^{m} x_i\right)^2},$$
$$a_1 = \frac{m \sum_{i=1}^{m} x_i y_i - \sum_{i=1}^{m} x_i \sum_{i=1}^{m} y_i}{m\left(\sum_{i=1}^{m} x_i^2\right) - \left(\sum_{i=1}^{m} x_i\right)^2}.$$

Polynomial Least Squares

- Use the least squares approach for approximating a set of data, $\{(x_i, y_i) \mid i = 1, 2, \ldots, m\}$, with a polynomial

$$P_n(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0,$$

  of degree $n < m - 1$.

- Choose $a_0, a_1, \ldots, a_n$ to minimize the least squares error $E = E_2(a_0, \ldots, a_n)$, where

$$E = \sum_{i=1}^{m} (y_i - P_n(x_i))^2.$$

Polynomial Least Squares (cont'd)

- We want $\frac{\partial E}{\partial a_j} = 0$ for each $j = 0, 1, \ldots, n$, and this gives $n+1$ normal equations:

$$
\begin{aligned}
a_0 \sum_{i=1}^m x_i^0 + a_1 \sum_{i=1}^m x_i^1 + a_2 \sum_{i=1}^m x_i^2 + \cdots + a_n \sum_{i=1}^m x_i^n &= \sum_{i=1}^m y_i x_i^0, \\
a_0 \sum_{i=1}^m x_i^1 + a_1 \sum_{i=1}^m x_i^2 + a_2 \sum_{i=1}^m x_i^3 + \cdots + a_n \sum_{i=1}^m x_i^{n+1} &= \sum_{i=1}^m y_i x_i^1, \\
&\vdots \\
a_0 \sum_{i=1}^m x_i^n + a_1 \sum_{i=1}^m x_i^{n+1} + a_2 \sum_{i=1}^m x_i^2 + \cdots + a_n \sum_{i=1}^m x_i^{2n} &= \sum_{i=1}^m y_i x_i^n.
\end{aligned}
$$

- This has a unique solution when x_i are distinct.

Other Least Squares

- When the data are exponentially related (rather than linear), one should consider the form

$$y = be^{ax} \quad \text{or} \quad y = bx^a$$

for some constants a and b.

- This requires minimizing

$$E = \sum_{i=1}^{m} (y_i - be^{ax_i})^2 \quad \text{or} \quad E = \sum_{i=1}^{m} (y_i - bx_i^a)^2,$$

which in general do not have a closed-form solution for $\frac{\partial E}{\partial a} = \frac{\partial E}{\partial b} = 0$.

Q. How can we circumvent such issue?

Other Least Squares (cont'd)

- Instead, one could consider

$$\ln y = \ln b + ax \quad \text{or} \quad \ln y = \ln b + a \ln x.$$

- Then we minimize

$$E = \sum_{i=1}^{m} [\ln y_i - (\ln b + ax)]^2 \quad \text{or} \quad E = \sum_{i=1}^{m} [\ln y_i - (\ln b + a \ln x)]^2.$$

Q. Do they produce results the same as the original least squares below?

$$E = \sum_{i=1}^{m} (y_i - be^{ax})^2 \quad \text{or} \quad E = \sum_{i=1}^{m} (y_i - bx^a)^2$$

Least Squares Approximation

- Suppose $f \in C[a,b]$, and we want to approximate it by a polynomial

$$P_n(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 = \sum_{k=0}^{n} a_k x^k$$

that minimizes the error

$$\begin{aligned} E &\equiv E_2(a_0, a_1, \ldots, a_n) \\ &= \int_a^b [f(x) - P_n(x)]^2 dx = \int_a^b \left(f(x) - \sum_{k=0}^{n} a_k x^k \right)^2 dx \\ &= \end{aligned}$$

Least Squares Approximation (cont'd)

- A necessary condition for $a_0, a_1, \ldots, a_n$ to minimize E is that

  for each $j = 0, 1, \ldots, n$.

- Hence, to find $P_n(x)$, the $(n+1)$ normal equations

$$\sum_{k=0}^{n} a_k \int_a^b x^{j+k} dx = \int_a^b x^j f(x) dx, \quad \text{for each } j = 0, 1, \ldots, n,$$

  must be solved for the $(n+1)$ unknowns a_j.

- These equations always have a unique solution when $f \in C[a, b]$.

Least Squares Approximation (cont'd)

Ex. Find the least squares approximating polynomial of degree 2 for the function $f(x) = \sin \pi x$ on the interval $[0, 1]$.

Sol. The normal equations for $P_2(x) = a_2 x^2 + a_1 x + a_0$ are

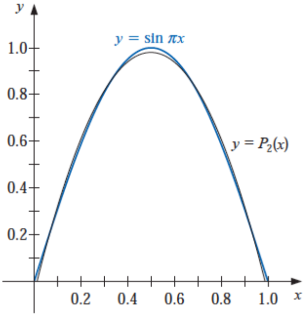
Least Squares Approximation (cont'd)

- Performing the integration yields

$$a_0 + \frac{1}{2}a_1 + \frac{1}{3}a_2 = \frac{2}{\pi}, \quad \frac{1}{2}a_0 + \frac{1}{3}a_1 + \frac{1}{4}a_2 = \frac{1}{\pi}, \quad \frac{1}{3}a_0 + \frac{1}{4}a_1 + \frac{1}{5}a_2 = \frac{\pi^2 - 4}{\pi^3}.$$

- The solution is

$$a_0 = \frac{12\pi^2 - 120}{\pi^3} \approx -0.050465, \quad a_1 = -a_2 = \frac{720 - 60\pi^2}{\pi^3} \approx 4.12251.$$



Least Squares Approximation (cont'd)

- The coefficients in the linear system are of the form

$$\int_a^b x^{j+k} dx = \frac{b^{j+k+1} - a^{j+k+1}}{j+k+1}$$

  which is known as **Hilbert matrix**.

- This is a classic example for demonstrating round-off error difficulties.

Least Squares Approximation (cont'd)

Ex. The $n \times n$ Hilbert matrix $H^{(n)}$ defined by

$$H_{ij}^{(n)} = \frac{1}{i+j-1}, \quad 1 \leq i, j \leq n,$$

is an ill-conditioned matrix. Compute $K_\infty(H^{(n)})$ for $n = 4, 5$.

Sol. $K_\infty(H^{(4)}) = 28375$ and $K_\infty(H^{(5)}) = 943656$.

Linearly Independent Functions

- Let's study a computationally more efficient approach, which is easy to determine $P_{n+1}(x)$, once $P_n(x)$ is known. (The previous method and the Lagrange polynomial are not computationally efficient in that sense.)

Definition 1

The set of functions $\{\phi_0, \ldots, \phi_n\}$ *is said to be* **linearly independent** *on* $[a, b]$ *if, whenever*

$$c_0\phi_0(x) + c_1\phi_1(x) + \cdots + c_n\phi_n(x) = 0, \quad \textit{for all } x \in [a, b],$$

we have $c_0 = c_1 = \cdots = c_n = 0$. *Otherwise, the set of functions is said to be* **linearly dependent**.

Linearly Independent Functions (cont'd)

Theorem 1

Suppose that, for each $j = 0, 1, \ldots, n$, $\phi_j(x)$ is a polynomial of degree j. Then $\{\phi_0, \ldots, \phi_n\}$ is linearly independent on any interval $[a, b]$.

- Let $\prod_n$ denote the set of all polynomials degree at most n.

Theorem 2

Suppose that $\{\phi_0(x), \phi_1(x), \ldots, \phi_n(x)\}$ is a collection of linearly independent polynomials in $\prod_n$. Then any polynomial in $\prod_n$ can be written uniquely as a linear combination of $\phi_0(x), \phi_1(x), \ldots, \phi_n(x)$.

Linearly Independent Functions (cont'd)

- Suppose $\{\phi_0, \phi_1, \ldots, \phi_n\}$ is a set of linearly independent functions on $[a, b]$. Given $f \in C[a, b]$, we seek a linear combination

$$P(x) = \sum_{k=0}^{n} a_k \phi_k(x)$$

to minimize the error

$$E = E(a_0, \ldots, a_n) = \int_a^b \left[f(x) - \sum_{k=0}^{n} a_k \phi_k(x) \right]^2 dx.$$

- We further need to introduce the notions of weight functions and orthogonality. Why?

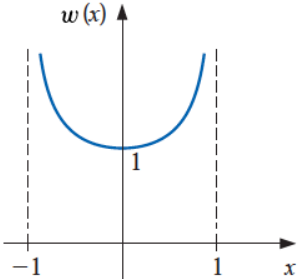
Weight Functions

- A weight function assigns varying degrees of importance to approximations on certain portions of the interval.

- For example, the weight function

$$w(x) = \frac{1}{\sqrt{1 - x^2}}$$

places less emphasis near the center of the interval $(-1, 1)$ and more emphasis when $|x|$ is near 1.



Orthogonal Functions

- Suppose $\{\phi_0, \phi_1, \ldots, \phi_n\}$ is a set of linearly independent functions on $[a, b]$ and w is a weight function for $[a, b]$. Given $f \in C[a, b]$, we seek a linear combination

$$P(x) = \sum_{k=0}^{n} a_k \phi_k(x)$$

to minimize the error

$$E = E(a_0, \ldots, a_n) = \int_a^b w(x) \left[f(x) - \sum_{k=0}^{n} a_k \phi_k(x) \right]^2 dx.$$

Orthogonal Functions (cont'd)

- We want for each $j = 0, 1, \ldots, n$,

$$0 = \frac{\partial E}{\partial a_j} = 2 \int_a^b w(x) \left[f(x) - \sum_{k=0}^{n} a_k \phi_k(x) \right] \phi_j(x) dx,$$

which gives the system of normal equations

$$\int_a^b w(x) f(x) \phi_j(x) dx = \sum_{k=0}^{n} a_k \int_a^b w(x) \phi_k(x) \phi_j(x) dx, \quad \text{for } j = 0, 1, \ldots, n.$$

Q. How should we choose $\phi_k(x)$?

Orthogonal Functions (cont'd)

- If the functions $\phi_0, \phi_1, \ldots, \phi_n$ is chosen so that

$$\int_a^b w(x)\phi_k(x)\phi_j(x)dx = \begin{cases} 0, & \text{when } j \neq k, \\ \alpha_j, & \text{when } j = k, \end{cases}$$

for positive constants α_j, the normal equations reduces to

$$\int_a^b w(x)f(x)\phi_j(x)dx = a_j\alpha_j,$$

for each $j = 0, 1, \ldots, n$.

Orthogonal Functions (cont'd)

Definition 2

$\{\phi_0, \phi_1, \ldots, \phi_n\}$ *is said to be an* **orthogonal set of functions** *for the interval* $[a, b]$ *with respect to the weight function* w *if*

$$\int_a^b w(x)\phi_k(x)\phi_j(x)dx = \begin{cases} 0, & \text{when } j \neq k, \\ \alpha_j > 0, & \text{when } j = k. \end{cases}$$

If, in addition, $\alpha_j = 1$ *for each* $j = 0, 1, \ldots, n$*, the set is said to be* **orthonormal**.

Orthogonal Functions (cont'd)

Theorem 3

If $\{\phi_0, \ldots, \phi_n\}$ *is an orthogonal set of functions on an interval* $[a, b]$ *with respect to the weight function* w*, then the least squares approximation to* f *on* $[a, b]$ *with respect to* w *is*

$$P(x) = \sum_{j=0}^{n} a_j \phi_j(x),$$

where, for each $j = 0, 1, \ldots, n$*,*

$$a_j = \frac{\int_a^b w(x)\phi_j(x) f(x) dx}{\int_a^b w(x)[\phi_j(x)]^2 dx} = \frac{1}{\alpha_j} \int_a^b w(x)\phi_j(x) f(x) dx.$$

- We next particularly focus on the orthogonal sets of **polynomials**, and the theorem next, based on the **Gram-Schmidt process**, describes how to construct orthogonal polynomials on $[a, b]$ with respect to w.

Orthogonal Polynomials

Theorem 4

The set of polynomial functions $\{\phi_0, \phi_1, \ldots, \phi_n\}$ *defined in the following way is orthogonal on* $[a, b]$ *with respect to the weight function* w*:*

$$\phi_0(x) = 1, \quad \phi_1(x) = x - B_1, \quad \text{for each } x \text{ in } [a, b],$$

where

$$B_1 = \frac{\int_a^b x w(x)[\phi_0(x)]^2 dx}{\int_a^b w(x)[\phi_0(x)]^2 dx},$$

and when $k \geq 2$*,*

$$\phi_k(x) = (x - B_k)\phi_{k-1}(x) - C_k \phi_{k-2}(x), \quad \text{for each } x \text{ in } [a, b],$$

where

$$B_k = \frac{\int_a^b x w(x)[\phi_{k-1}(x)]^2 dx}{\int_a^b w(x)[\phi_{k-1}(x)]^2 dx} \quad \text{and} \quad C_k = \frac{\int_a^b x w(x)\phi_{k-1}(x)\phi_{k-2}(x) dx}{\int_a^b w(x)[\phi_{k-2}(x)]^2 dx}.$$

Orthogonal Polynomials (cont'd)

Corollary 1

For any $n > 0$, the set of polynomial functions $\{\phi_0, \ldots, \phi_n\}$ given in Theorem 4 is linearly independent of $[a, b]$ and

$$\int_a^b w(x)\phi_n(x)Q_k(x)dx = 0,$$

for any polynomial $Q_k(x)$ of degree $k < n$.

Proof For each $k = 0, 1, \ldots, n$, $\phi_k(x)$ is a polynomial of degree k. So Theorem 1 implies that $\{\phi_0, \ldots, \phi_n\}$ is a linearly independent set. By Theorem 2 there exist numbers $c_0, \ldots, c_k$ such that

$$Q_k(x) = \sum_{j=0}^{k} c_j\phi_j(x).$$

Because ϕ_n is orthogonal to ϕ_j for each $j = 0, 1, \ldots, k$ we have

$$\int_a^b w(x)Q_k(x)\phi_n(x) = \sum_{j=0}^{k} c_j \int_a^b w(x)\phi_j(x)\phi_n(x)dx = 0.$$

Orthogonal Polynomials (cont'd)

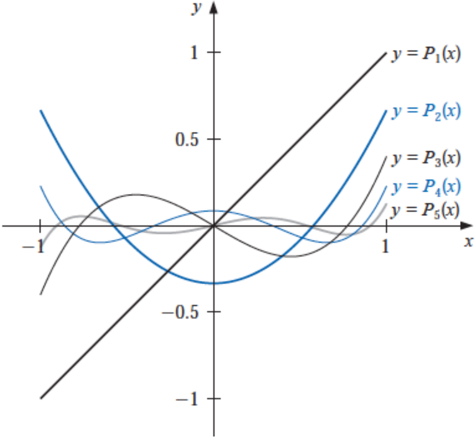
- Recall: In Section 4.7, the roots of Legendre polynomials are used as the nodes in Gaussian quadrature.
- The set of **Legendre polynomials** is orthogonal on $[-1, 1]$ with respect to the weight function $w(x) = 1$. (Note that the classical definition of the Legendre polynomials requires that $P_n(1) = 1$ for each n.)
- Using the Gram-Schmidt process with $P_0(x) = 1$ gives

$$B_1 = \frac{\int_{-1}^{1} x dx}{\int_{-1}^{1} 1 dx} = 0 \quad \text{and} \quad P_1(x) = x - B_1 = x.$$

Also,

Orthogonal Polynomials (cont'd)

- The higher-degree Legendre polynomials are shown below.



Chebyshev Polynomials

- The Chebyshev polynomials $\{T_n(x)\}$ for $x \in [-1,1]$, defined by

$$T_n(x) = \cos(n \arccos x) \quad \text{for each } n \geq 0,$$

are orthogonal on $(-1,1)$ with respect to the weight function

$$w(x) = \frac{1}{\sqrt{1-x^2}}.$$

Q. Why such weight function?

Q. Is $\{T_n(x)\}$ a (orthogonal) set of polynomials?

Chebyshev Polynomials (cont'd)

- The Chebyshev polynomials are used to minimize approximation error. We will soon study how they are used to solve two problems of this type:

1. an optimal placing of interpolating points to minimize the error in Lagrange interpolation
2. a means of reducing the degree of an approximating polynomial with minimal loss of accuracy.

Chebyshev Polynomials (cont'd)

- For $n = 0, 1$ we have

$$T_0(x) = \cos 0 = 1 \quad \text{and} \quad T_1(x) = \cos(\arccos x) = x.$$

- For $n \geq 1$, introducing $\theta = \arccos x$ yields

$$T_n(\theta(x)) \equiv T_n(\theta) = \cos(n\theta), \quad \text{where } \theta \in [0, \pi].$$

We then have

$$\begin{aligned} T_{n+1}(\theta) &= \cos((n+1)\,\theta) = \cos\theta\cos(n\theta) - \sin\theta\sin(n\theta) \\ T_{n-1}(\theta) &= \cos((n-1)\,\theta) = \cos\theta\cos(n\theta) + \sin\theta\sin(n\theta)\,. \end{aligned}$$

Chebyshev Polynomials (cont'd)

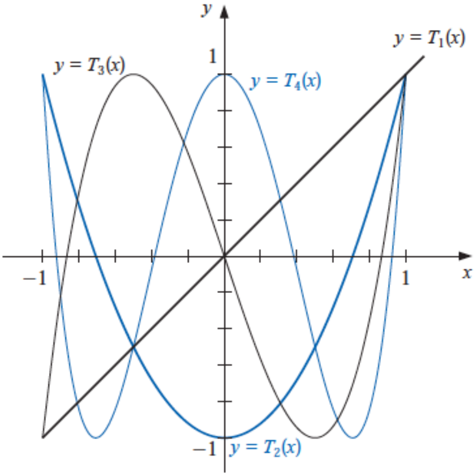
- Since $T_0(x) = 1$ and $T_1(x) = x$, we have

$$\begin{aligned} T_2(x) &= 2xT_1(x) - T_0(x) = 2x^2 - 1, \\ T_3(x) &= 2xT_2(x) - T_1(x) = 4x^3 - 3x, \\ T_4(x) &= 2xT_3(x) - T_2(x) = 8x^4 - 8x^2 + 1, \\ &\vdots \end{aligned}$$

- When $n \geq 1$, $T_n(x)$ is a polynomial of degree n with leading coefficient 2^{n-1}.

Chebyshev Polynomials (cont'd)

- The graphs of Chebyshev polynomials are shown below.



Chebyshev Polynomials (cont'd)

- Let's show the orthogonality of the Chebyshev polynomials with respect to $w(x) = (1 - x^2)^{-1/2}$, using

$$\int_{-1}^{1} \frac{T_n(x)T_m(x)}{\sqrt{1-x^2}} dx = \int_{-1}^{1} \frac{\cos(n \arccos x)\cos(m \arccos x)}{\sqrt{1-x^2}} dx$$

- Using $\theta = \arccos x$ gives

$$d\theta = -\frac{1}{\sqrt{1-x^2}} dx$$

and thus

$$\int_{-1}^{1} \frac{T_n(x)T_m(x)}{\sqrt{1-x^2}} dx =$$

Chebyshev Polynomials (cont'd)

- Similarly, we can show that for each $n \geq 1$,

$$\int_{-1}^{1} \frac{[T_n(x)]^2}{\sqrt{1-x^2}} dx = \frac{\pi}{2}.$$

Chebyshev Polynomials (cont'd)

- Recall: The Chebyshev polynomials are used to minimize approximation error. We will next study how they are used to solve two problems of this type:

1. an optimal placing of interpolating points to minimize the error in Lagrange interpolation
2. a means of reducing the degree of an approximating polynomial with minimal loss of accuracy.

- Then, what should we study next?

Chebyshev Polynomials (cont'd)

Theorem 5

The Chebyshev polynomial $T_n(x)$ of degree $n \geq 1$ has n simple zeros in $[-1, 1]$ at

$$\bar{x}_k = \cos\left(\frac{2k-1}{2n}\pi\right), \quad \text{for each } k = 1, 2, \ldots, n.$$

Moreover, $T_n(x)$ has its absolute extrema at

$$\bar{x}'_k = \cos\left(\frac{k\pi}{n}\right) \quad \text{with} \quad T_n(\bar{x}'_k) = (-1)^k, \quad \text{for each } k = 0, 1, \ldots, n.$$

Proof We have

$$T_n(\bar{x}_k) = \cos(n \arccos \bar{x}_k) = \cos\left(\frac{2k-1}{2}\pi\right) = 0.$$

Since $\bar{x}_k$ are distinct (since $\cos$ is a strictly decreasing function from 0 to π) and $T_n(x)$ is a polynomial of degree n, so all the zero of $T_n(x)$ must have this form.

Monic Chebyshev Polynomials

- The monic (polynomials with leading coefficient 1) Chebyshev polynomials $\tilde{T}_n(x)$ are derived by dividing by 2^{n-1} as

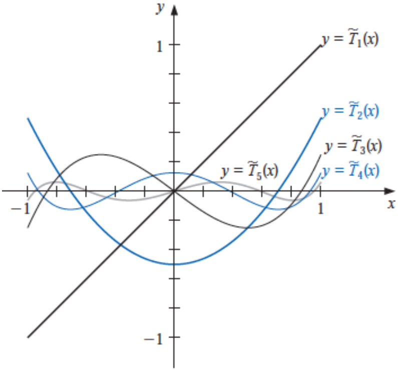
$$\tilde{T}_0(x) = 1 \quad \text{and} \quad \tilde{T}_n(x) = \frac{1}{2^{n-1}} T_n(x), \quad \text{for each } n \geq 1.$$

- They have the following relationship

$$\begin{aligned} \tilde{T}_2(x) &= x\tilde{T}_1(x) - \frac{1}{2}\tilde{T}_0(x) \\ \tilde{T}_{n+1} &= x\tilde{T}_n(x) - \frac{1}{4}\tilde{T}_{n-1}(x), \quad \text{for each } n \geq 2. \end{aligned}$$

Monic Chebyshev Polynomials (cont'd)

- The graphs of monic Chebyshev polynomials are shown below.



- The extreme values of $\tilde{T}_n(x)$ for $n \geq 1$, occur at

$$\bar{x}'_k = \cos\left(\frac{k\pi}{n}\right), \quad \text{with } \tilde{T}_n(\bar{x}'_k) = \frac{(-1)^k}{2^{n-1}}, \quad \text{for each } k = 0, 1, 2, \ldots, n.$$

Monic Chebyshev Polynomials (cont'd)

- Let $\tilde{\prod}_n$ denote the set of all monic polynomials of degree n.

Theorem 6

The polynomials of the form $\tilde{T}_n(x)$, when $n \geq 1$, have the property that

$$\frac{1}{2^{n-1}} = \max_{x\in[-1,1]} |\tilde{T}_n(x)| \leq \max_{x\in[-1,1]} |P_n(x)|, \quad \textit{for all } P_n(x) \in \tilde{\prod}_n.$$

Moreover, equality holds only if $P_n = \tilde{T}_n$.

Proof Suppose that $P_n(x) \in \tilde{\prod}_n$ and that

$$\max_{x\in[-1,1]} |P_n(x)| < \frac{1}{2^{n-1}} = \max_{x\in[-1,1]} |\tilde{T}_n(x)|.$$

The Use of Chebyshev Polynomials

- Minimizing Lagrange Interpolation Error on $[-1\ 1]$ (and on Arbitrary Intervals)
- Reducing the Degree of Approximating Polynomials

Minimizing Lagrange Interpolation Error

Q. Where should we place interpolating nodes to minimize the error in Lagrange interpolation?

- Recall:

Theorem 7

Suppose $x_0, x_1, \ldots, x_n$ are distinct numbers in the interval $[a, b]$ and $f \in C^{n+1}[a, b]$. Then, for each x in $[a, b]$, a number $\xi(x)$ (generally unknown) between $x_0, x_1, \ldots, x_n$, and hence in (a, b), exists with

$$f(x) = P(x) + \frac{f^{(n+1)}(\xi(x))}{(n+1)!}(x - x_0)(x - x_1)\cdots(x - x_n)$$

where $P(x)$ is the nth Lagrange interpolating polynomial.

Minimizing Lagrange Interpolation Error (cont'd)

- In general, we cannot control $\xi(x)$, so we choose nodes to minimize

$$|(x-x_0)(x-x_1)\cdots(x-x_n)|$$

throughout the interval $[-1,1]$.

- Since $(x-x_0)(x-x_1)\cdots(x-x_n)$ is a monic polynomial of degree $(n+1)$, the minimum is obtained when it is

Minimizing Lagrange Interpolation Error (cont'd)

- In other words, $|(x-x_0)(x-x_1)\cdots(x-x_n)|$ is smallest when x_k is chosen to be the $(k+1)$th zero fo $\tilde{T}_{n+1}$; that is,

$$\bar{x}_{k+1} =$$

- Since $\max_{x\in[-1,1]} |\tilde{T}_{n+1}(x)| = 2^{-n}$, we have

$$\frac{1}{2^n} = \max_{x\in[-1,1]} |(x-\bar{x}_1)\cdots(x-\bar{x}_{n+1})| \leq \max_{x\in[-1,1]} |(x-x_0)\cdots(x-x_n)|,$$

  for any choice of $x_0, x_1, \ldots, x_n$ in the interval $[-1, 1]$.

Minimizing Lagrange Interpolation Error (cont'd)

Corollary 2

Suppose that $P(x)$ is the interpolating polynomial of degree at most n with nodes at the zeros of $T_{n+1}(x)$. Then

$$\max_{x\in[-1,1]} |f(x) - P(x)| \le \frac{1}{2^n(n+1)!} \max_{x\in[-1,1]} |f^{(n+1)}(x)|,$$

for each $f \in C^{n+1}[-1,1]$.

Minimizing Approximation Error on Arbitrary Intervals

- For a general closed interval $[a, b]$, use the change of variables

$$\tilde{x} = \frac{1}{2}[(b-a)x + a + b]$$

to transform the numbers $\bar{x}_k$ in $[-1, 1]$ into $\tilde{x}_k$ in $[a, b]$.

Ex. Let $f(x) = xe^x$ on $[0, 1.5]$. Compare the values given by the Lagrange polynomial with four equally spaced nodes and with nodes given by zeros of the fourth Chebyshev polynomial.

Sol. Equally spaced nodes $x_0 = 0$, $x_1 = 0.5$, $x_2 = 1$, $x_3 = 1.5$ gives

$$P_3(x) = 1.3875x^3 + 0.057570x^2 + 1.2730x.$$

Minimizing Approx. Error on Arbitrary Intervals (cont'd)

Sol. Zeros of the fourth Chebyshev polynomial are

$$\bar{x}_1 = \cos\frac{\pi}{8} = 0.92388, \quad \bar{x}_2 = \cos\frac{3\pi}{8} = 0.38268,$$
$$\bar{x}_3 = \cos\frac{5\pi}{8} = -0.38268, \quad \bar{x}_4 = \cos\frac{7\pi}{8} = -0.92388.$$

Then using the transformation

$$\tilde{x}_k = \frac{1}{2}[(1.5-0)\bar{x}_k + (1.5+0)] = 0.75 + 0.75\bar{x}_k,$$

we have

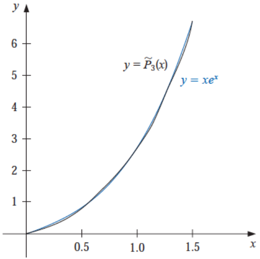
$$\tilde{x}_1 = 1.44291, \quad \tilde{x}_2 = 1.03701, \quad \tilde{x}_3 = 0.46299, \quad \tilde{x}_4 = 0.05709,$$

and this gives

$$\tilde{P}_3(x) = 1.3811x^3 + 0.044652x^2 + 1.3031x - 0.014352.$$

Minimizing Approx. Error on Arbitrary Intervals (cont'd)

x	$f(x) = xe^x$	$P_3(x)$	$\|xe^x - P_3(x)\|$	$\tilde{P}_3(x)$	$\|xe^x - \tilde{P}_3(x)\|$
0.15	0.1743	0.1969	0.0226	0.1868	0.0125
0.25	0.3210	0.3435	0.0225	0.3358	0.0148
0.35	0.4967	0.5121	0.0154	0.5064	0.0097
0.65	1.245	1.233	0.012	1.231	0.014
0.75	1.588	1.572	0.016	1.571	0.017
0.85	1.989	1.976	0.013	1.974	0.015
1.15	3.632	3.650	0.018	3.644	0.012
1.25	4.363	4.391	0.028	4.382	0.019
1.35	5.208	5.237	0.029	5.224	0.016



Reducing the Degree of Approximating Polynomials

- Chebyshev polynomials can also be used to reduce the degree of an approximating polynomial with a minimal loss of accuracy.

- Consider approximating an arbitrary nth-degree polynomial

$$P_n(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$$

  on $[-1, 1]$ with a polynomial of degree at most $n - 1$.

- The goal is to choose $P_{n-1}(x)$ in $\prod_{n-1}$ so that

$$\max_{x \in [-1,1]} |P_n(x) - P_{n-1}(x)|$$

  is as small as possible.

Reducing the Degree of Approximating Poly. (cont'd)

- Since $(P_n(x) - P_{n-1}(x))/a_n$ is a monic polynomial of degree n, we have

$$\max_{x\in[-1,1]} \left| \frac{1}{a_n}(P_n(x) - P_{n-1}(x)) \right| \geq \frac{1}{2^{n-1}}.$$

The equality holds when

$$\frac{1}{a_n}(P_n(x) - P_{n-1}(x)) = \tilde{T}_n(x).$$

- This means that we should choose

$$P_{n-1}(x) =$$

which satisfies

$$\max_{x\in[-1,1]} |P_n(x) - P_{n-1}(x)| = |a_n| \max_{x\in[-1,1]} \left| \frac{1}{a_n}(P_n(x) - P_{n-1}(x)) \right| = \frac{|a_n|}{2^{n-1}}.$$

Reducing the Degree of Approximating Poly. (cont'd)

Ex. Consider the fourth Taylor polynomial

$$P_4(x) = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24},$$

for $f(x) = e^x$ about $x_0 = 0$ on the interval $[-1, 1]$, with a truncation error

$$|f(x) - P_4(x)| = \frac{|f^{(5)}(\xi(x))|\,|x^5|}{120} \leq \frac{e}{120} \approx 0.023, \quad \text{for} \ -1 \leq x \leq 1.$$

Reduce the degree of the approximating polynomial within an error bound of 0.05.

Reducing the Degree of Approximating Poly. (cont'd)

Sol. The polynomial of degree 3 or less that best uniformly approximates $P_4(x)$ on $[-1, 1]$ is

$$P_3(x) =$$

Reducing the Degree of Approximating Poly. (cont'd)

Sol. The polynomial of degree 2 or less that best uniformly approximates $P_3(x)$ on $[-1, 1]$ is

$$\begin{aligned} P_2(x) = P_3(x) - \frac{1}{6}\tilde{T}_3(x) &= \frac{1}{6}x^3 \frac{13}{24}x^2 + \frac{191}{192} - \frac{1}{6}\left(x^3 - \frac{3}{4}x\right) \\ &= \frac{13}{24}x^2 + \frac{9}{8}x + \frac{191}{192}, \end{aligned}$$

which has the error

$$|P_3(x) - P_2(x)| = \left|\frac{1}{6}\tilde{T}_3(x)\right| \le \frac{1}{2^2 6} = \frac{1}{24} \approx 0.042.$$

The total error is

$$\begin{aligned} |f(x) - P_2(x)| &\le |f(x) - P_4(x)| + |P_4(x) - P_3(x)| + |P_3(x) - P_2(x)| \\ &\le 0.023 + 0.0053 + 0.042 = 0.0703, \end{aligned}$$

which exceeds the tolerance of 0.05.

Reducing the Degree of Approximating Poly. (cont'd)

x	e^x	$P_4(x)$	$P_3(x)$	$P_2(x)$	$\lvert e^x - P_2(x)\rvert$
-0.75	0.47237	0.47412	0.47917	0.45573	0.01664
-0.25	0.77880	0.77881	0.77604	0.74740	0.03140
0.00	1.00000	1.00000	0.99479	0.99479	0.00521
0.25	1.28403	1.28402	1.28125	1.30990	0.02587
0.75	2.11700	2.11475	2.11979	2.14323	0.02623