Chapter 10

Orthogonal Polynomials and Least-Squares Approximations to Functions

**4/5/13 ET

10.1 Discrete Least-Squares Approximations

Given a set of data points (x_1, y_1), (x_2, y_2), ..., (x_m, y_m), a normal and useful practice in many applications in statistics, engineering and other applied sciences is to construct a curve that is considered to be the "best fit" for the data, in some sense.

Several types of "fits" can be considered. But the one that is used most in applications is the "**least-squares fit**". Mathematically, the problem is the following:

Discrete Least-Squares Approximation Problem

Given a set of n discrete data points (x_i, y_i), $i = 1, 2, \ldots, m$.

Find the algebraic polynomial

$$P_n(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n (n < m)$$

such that the error $E(a_0, a_1, \ldots, a_n)$ in the least-squares sense in minimized; that is,

$$E(a_0, a_1, \ldots, a_n) = \sum_{i=1}^{m} (y_i - a_0 - a_1 x_i - a_2 x_i^2 - \cdots - a_n x_i^n)^2$$

is minimum.

Here $E(a_0, a_1, \dots, a_n)$ is a function of $(n+1)$ variables: $a_0, a_1, \dots, a_n$.

Since $E(a_0, a_1, \dots, a_n)$ is a function of the variables, $a_0, a_1, \dots, a_n$, for this function to be minimum, we must have:

$$\frac{\partial E}{\partial a_j} = 0, \quad j = 0, 1, \dots, n.$$

Now, simple computations of these partial derivatives yield:

$$\begin{aligned}
\frac{\partial E}{\partial a_0} &= -2\sum_{i=1}^{m}(y_i - a_0 - a_1 x_i - \cdots - a_n x_i^n) \\
\frac{\partial E}{\partial a_1} &= -2\sum_{i=1}^{m} x_i(y_i - a_0 - a_1 x_i - \cdots - a_n x_i^n) \\
&\vdots \\
\frac{\partial E}{\partial a_n} &= -2\sum_{i=1}^{m} x_i^n(y_i - a_0 - a_1 x_i - \cdots - a_n x_i^n)
\end{aligned}$$

Setting these equations to be zero, we have

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

Set

$$\begin{aligned}
s_k &= \sum_{i=1}^{m} x_i^k, \quad k = 0, 1, \dots, 2n \\
b_k &= \sum_{i=1}^{m} x_i^k y_i, \quad k = 0, 1, \dots, n
\end{aligned}$$

Using these notations, the above equations can be written as:

$$
\begin{aligned}
&s_0a_0 + s_1a_1 + \cdots + s_na_n = b_0 \quad \left(\text{Note that } \sum_{i=1}^{m} 1 = \sum_{i=1}^{m} x_i^0 = s_0.\right) \\
&s_1a_0 + s_2a_1 + \cdots + s_{n+1}a_n = b_1 \\
&\vdots \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \vdots \\
&s_na_0 + s_{n+1}a_1 + \cdots + s_{2n}a_n = b_n
\end{aligned}
$$

This is a system of $(n+1)$ equations in $(n+1)$ unknowns $a_0, a_1, \ldots, a_n$. These equations are called **Normal Equations**. This system now can be solved to obtain these $(n+1)$ unknowns, provided a solution to the system exists. We will not *show that this system has a unique solution if x_i's are distinct.*

The system can be written in the following matrix form:

$$
\begin{pmatrix} s_0 & s_1 & \cdots & s_n \\ s_1 & s_2 & \cdots & s_{n+1} \\ \vdots & \vdots & & \vdots \\ s_n & s_{n+1} & \cdots & s_{2n} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_n \end{pmatrix}.
$$

or

$$Sa = b$$

where

$$
S = \begin{pmatrix} s_0 & s_1 & \cdots & s_n \\ s_1 & s_2 & \cdots & s_{n+1} \\ \vdots & \vdots & & \vdots \\ s_n & s_{n+1} & \cdots & s_{2n} \end{pmatrix}, \quad a = \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{pmatrix}, \quad b = \begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_n \end{pmatrix}.
$$

Define

$$
V = \begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^n \\ 1 & x_2 & x_2^2 & \cdots & x_2^n \\ 1 & x_3 & x_3^2 & \cdots & x_3^n \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_m & x_m^2 & \cdots & x_m^n \end{pmatrix}
$$

Then the above system has the form:

$$V^T V a = b.$$

The matrix V is known as the **Vandermonde matrix**, and we have seen in Chapter 6 that *this matrix has full rank if x_i's are distinct.* In this case, the matrix $S = V^T V$ is symmetric and positive definite [**Exercise**] and is therefore nonsingular. Thus, if x_i's are distinct, the equation $Sa = b$ has a unique solution.

Theorem 10.1 (Existence and uniqueness of Discrete Least-Squares Solutions). Let (x_1, y_1), (x_2, y_2), …, (x_n, y_n) be n distinct points. Then the discrete least-square approximation problem has a unique solution.

10.1.1 Least-Squares Approximation of a Function

We have described least-squares approximation to fit a set of discrete data. Here we describe continuous least-square approximations of a function $f(x)$ by using polynomials.

First, consider approximation by a polynomial with **monomial basis**: $\{1, x, x^2, \ldots, x^n\}$.

Least-Square Approximations of a Function Using Monomial Polynomials

Given a function $f(x)$, continuous on $[a, b]$, find a polynomial $P_n(x)$ of degree at most n:

$$P_n(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$$

such that the integral of the square of the error is minimized. That is,

$$E = \int_a^b [f(x) - P_n(x)]^2 \, dx$$

is minimized.

The polynomial $P_n(x)$ is called the **Least-Squares Polynomial**.

Since E is a function of $a_0, a_1, \ldots, a_n$, we denote this by $E(a_0, a_1, \ldots, a_n)$.

For minimization, we must have

$$\frac{\partial E}{\partial a_i} = 0, \quad i = 0, 1, \ldots, n.$$

As before, these conditions will give rise to a system of $(n + 1)$ **normal** equations in $(n + 1)$ unknowns: $a_0, a_1, \ldots, a_n$. Solution of these equations will yield the unknowns: $a_0, a_1, \ldots, a_n$.

10.1.2 Setting up the Normal Equations

Since

$$\begin{aligned}
E &= \int_a^b [f(x) - (a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n)]^2\,dx,\\
\frac{\partial E}{\partial a_0} &= -2\int_a^b [f(x) - a_0 - a_1 x - a_2 x^2 - \cdots - a_n x^n]\,dx\\
\frac{\partial E}{\partial a_1} &= -2\int_a^b x[f(x) - a_0 - a_1 x - a_2 x^2 - \cdots - a_n x^n]\,dx\\
&\ \ \vdots\\
\frac{\partial E}{\partial a_n} &= -2\int_a^b x^n[f(x) - a_0 - a_1 x - a_2 x^2 - \cdots - a_n x^n]\,dx
\end{aligned}$$

So,

$$\frac{\partial E}{\partial a_0} = 0 \Rightarrow a_0\int_a^b 1\,dx + a_1\int_a^b x\,dx + a_2\int_a^b x^2\,dx + \cdots + a_n\int_a^b x^n\,dx = \int_a^b f(x)\,dx$$

Similarly,

$$\frac{\partial E}{\partial a_i} = 0 \Rightarrow a_0\int_a^b x^i\,dx + a_1\int_a^b x^{i+1}\,dx + a_2\int_a^b x^{i+2}\,dx + \cdots + a_n\int_a^b x^{i+n}\,dx = \int_a^b x^i f(x)\,dx,$$
$$i = 1, 2, 3, \ldots, n.$$

So, the $(n+1)$ *normal equations* in this case are:

$$\begin{aligned}
i=0\colon\quad & a_0\int_a^b 1\,dx + a_1\int_a^b x\,dx + a_2\int_a^b x^2\,dx + \cdots + a_n\int_a^b x^n\,dx = \int_a^b f(x)\,dx\\
i=1\colon\quad & a_0\int_a^b x\,dx + a_1\int_a^b x^2\,dx + a_3\int_a^b x^3\,dx + \cdots + a_n\int_a^b x^{n+1}\,dx = \int_a^b x\,f(x)\,dx\\
& \vdots\\
i=n\colon\quad & a_0\int_a^b x^n\,dx + a_1\int_a^b x^{n+1}\,dx + a_2\int_a^b x^{n+2}\,dx + \cdots + a_n\int_a^b x^{2n}\,dx = \int_a^b x^n\,f(x)\,dx
\end{aligned}$$

Denote

$$\int_a^b x^i\,dx = s_i,\ i = 0, 1, 2, 3, \ldots, 2n,\ \text{and } b_i = \int_a^b x^i f(x)\,dx,\ i = 0, 1, \ldots, n.$$

Then the above $(n+1)$ equations can be written as

$$\begin{aligned}
a_0 s_0 + a_1 s_1 + a_2 s_2 + \cdots + a_n s_n &= b_0\\
a_0 s_1 + a_1 s_2 + a_2 s_3 + \cdots + a_n s_{n+1} &= b_1\\
&\ \vdots\\
a_0 s_n + a_1 s_{n+1} + a_2 s_{n+2} + \cdots + a_n s_{2n} &= b_n.
\end{aligned}$$

or in matrix notation

$$\begin{pmatrix} s_0 & s_1 & s_2 & \cdots & s_n \\ s_1 & s_2 & s_3 & \cdots & s_{n+1} \\ \vdots & \vdots & \vdots & & \vdots \\ s_n & s_{n+1} & s_{n+2} & \cdots & s_{2n} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_n \end{pmatrix}$$

Denote

$$S = (s_i), \quad a = \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{pmatrix}, \quad b = \begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_n \end{pmatrix}.$$

Then we have the system of normal equations:

$$Sa = b$$

The solution of these equations will yield the coefficients $a_0, a_1, \ldots, a_n$ *of the least-squares polynomial* $P_n(x)$.

A Special Case: Let the interval be $[0, 1]$. Then

$$s_i = \int_0^1 x^i \, dx = \frac{1}{i+1}, \quad i = 0, 1, \ldots, 2n.$$

Thus, in this case the matrix of the normal equations

$$S = \begin{pmatrix} 1 & \frac{1}{2} & \cdots & \frac{1}{n} \\ \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{n+1} \\ \vdots & \vdots & & \vdots \\ \frac{1}{n} & \frac{1}{n+1} & \cdots & \frac{1}{2n} \end{pmatrix}$$

which is a **Hilbert Matrix**. It is well-known to be **ill-conditioned** (see **Chapter 3**).

Algorithm 10.2 (Least-Squares Approximation using Monomial Polynomials).

Inputs: (i) $f(x)$ - A continuous function on $[a, b]$.

(ii) n - The degree of the desired least-squares polynomial

Output: The coefficients $a_0, a_1, \ldots, a_n$ of the desired least-squares polynomial: $P_n(x) = a_0 + a_1 x + \cdots + a_n x^n$.

Step 1. Compute $s_0, s_1, \ldots, s_{2n}$:
For $i = 0, 1, \ldots, 2n$ do
$s_i = \int_a^b x^i \, dx$
End

Step 2. Compute $b_0, b_1, \ldots, b_n$:
For $i = 0, 1, \ldots, n$ do
$b_i = \int_a^b x^i f(x) \, dx$
End

Step 3. Form the matrix S from the numbers $s_0, s_1, \ldots, s_{2n}$ and the vector b from the numbers $b_0, b_1, \ldots, b_n$.

$$S = \begin{pmatrix} s_0 & s_1 & \ldots & s_n \\ s_1 & s_2 & \ldots & s_{n+1} \\ \vdots & \vdots & & \vdots \\ s_n & s_{n+1} & \ldots & s_{2n} \end{pmatrix}, \quad b = \begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_n \end{pmatrix}.$$

Step 4. Solve the $(n+1) \times (n+1)$ system of equations for $a_0, a_1, \ldots a_n$:

$$Sa = b, \quad \text{where } a = \begin{pmatrix} a_0 \\ \vdots \\ a_n \end{pmatrix}.$$

Example 10.3

Find Linear and Quadratic least-squares approximations to $f(x) = e^x$ on $[-1, 1]$.

Linear Approximation: $n = 1$; $P_1(x) = a_0 + a_1 x$

$$s_0 = \int_{-1}^{1} dx = 2$$

$$s_1 = \int_{-1}^{1} x\,dx = \left[\frac{x^2}{2}\right]_{-1}^{1} = \frac{1}{2} - \left(\frac{1}{2}\right) = 0$$

$$s_2 = \int_{-1}^{1} x^2\,dx = \left[\frac{x^3}{3}\right]_{-1}^{1} = \frac{1}{3} - \left(\frac{-1}{3}\right) = \frac{2}{3}$$

Thus,

$$S = \begin{pmatrix} s_0 & s_1 \\ s_1 & s_2 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & \frac{2}{3} \end{pmatrix}$$

$$b_0 = \int_{-1}^{1} e^x\,dx = e - \frac{1}{e} = 2.3504$$

$$b_1 = \int_{-1}^{1} e^x x\,dx = \frac{2}{e} = 0.7358$$

The normal system is:

$$\begin{pmatrix} 2 & 0 \\ 0 & \frac{2}{3} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \begin{pmatrix} b_0 \\ b_1 \end{pmatrix}$$

This gives

$$a_0 = 1.1752, \quad a_1 = 1.1037$$

The linear least-squares polynomial $P_1(x) = 1.1752 + 1.1037x$.

Accuracy Check:

$$P_1(0.5) = 1.7270, \qquad e^{0.5} = 1.6487$$

Relative Error: $= 0.0453$.

$$\frac{|e^{0.5} - P_1(0.5)|}{|e^{0.5}|} = \frac{|1.6487 - 1.7270|}{|1.6487|} = 0.0475.$$

Quadratic Fitting: $n = 2$; $P_2(x) = a_0 + a_1 x + a_2 x^2$

$$s_0 = 2, \quad s_1 = 0, \quad s_2 = \frac{2}{3}$$

$$s_3 = \int_{-1}^{1} x^3 dx = \left[\frac{x^4}{4}\right]_{-1}^{1} = 0$$

$$s_4 = \int_{-1}^{1} x^4 dx = \left[\frac{x^5}{5}\right]_{-1}^{1} = \frac{2}{5}$$

$$
\begin{aligned}
b_0 &= \int_{-1}^{1} e^x dx = e - \frac{1}{e} = 2.3504 \\
b_1 &= \int_{-1}^{1} x e^x dx = \frac{2}{e} = 0.7358 \\
b_2 &= \int_{-1}^{1} x^2 e^x dx = e - \frac{5}{e} = 0.8789.
\end{aligned}
$$

The system of normal equations is:

$$
\begin{pmatrix} 2 & 0 & \frac{2}{3} \\ 0 & \frac{2}{3} & 0 \\ \frac{2}{3} & 0 & \frac{2}{5} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 2.3504 \\ 0.7358 \\ 0.8789 \end{pmatrix}
$$

The solution of this system is:

$$
a_0 = 0.9963, \quad a_1 = 1.1037, \quad a_2 = 0.5368.
$$

The quadratic least-squares polynomial $P_2(x) = 0.9963 + 1.1037x + 0.5368x^2$.

Accuracy Check:

$$
\begin{aligned}
P_2(0.5) &= 1.6889 \\
e^{0.5} &= 1.6487
\end{aligned}
$$

Relative error: $\frac{|P_2(0.5) - e^{0.5}|}{|e^{0.5}|} = \frac{|1.6824 - 1.6487|}{|1.6487|} = 0.0204.$

Example 10.4

Find linear and quadratic least-squares polynomial approximation to $f(x) = x^2 + 5x + 6$ in $[0, 1]$.

Linear Fit: $P_1(x) = a_0 + a_1 x$

$$
\begin{aligned}
s_0 &= \int_0^1 dx = 1 \\
s_1 &= \int_0^1 x dx = \frac{1}{2} \\
s_2 &= \int_0^1 x^2 dx = \frac{1}{3}
\end{aligned}
$$

$$\begin{aligned}
b_0 &= \int_0^1 (x^2+5x+6)\,dx = \frac{1}{3}+\frac{5}{2}+6 \\
&= \frac{53}{6} \\
b_1 &= \int_0^1 x(x^2+5x+6)\,dx = \int_0^1 (x^3+5x^2+6x)\,dx \\
&= \frac{1}{4}+\frac{5}{3}+\frac{6}{2} = \frac{59}{12}
\end{aligned}$$

The **normal equations** are:

$$\begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{3} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \begin{pmatrix} \frac{53}{6} \\ \frac{59}{12} \end{pmatrix} \Rightarrow \begin{matrix} a_0 = 5.8333 \\ a_1 = 6 \end{matrix}$$

The linear least squares polynomial $P_1(x) = 5.8333 + 6x$.

Accuracy Check:

Exact Value: $f(0.5) = 8.75$; $P_1(0.5) = 8.833$

Relative error: $\frac{|8.833-8.75|}{|8.75|} = 0.0095$.

Quadratic Least-Square Approximation: $P_2(x) = a_0 + a_1 x + a_2 x^2$

$$S = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{pmatrix}$$

$$\begin{aligned}
b_0 &= \frac{53}{6}, \quad b_1 = \frac{59}{12}, \\
b_2 &= \int_0^1 x^2(x^2+5x+6)\,dx = \int_0^1 (x^4+5x^3+6x^2)\,dx = \frac{1}{5}+\frac{5}{4}+\frac{6}{3} = \frac{69}{20}.
\end{aligned}$$

The solution of the linear system is: $a_0 = 6$, $a_1 = 5$, $a_2 = 1$,

$$P_2(x) = 6 + 5x + x^2 \ (\textbf{Exact})$$

10.1.3 Use of Orthogonal Polynomials in Least-squares Approximations

The least-squares approximation using monomial polynomials, as described above, is not numerically effective; since the system matrix S of normal equations is very often **ill-conditioned**.

For example, when the interval is $[0, 1]$, we have seen that S is a **Hilbert matrix**, which is notoriously ill-conditioned for even modest values of n. When $n = 5$, the condition number of this matrix $= \text{cond}(S) = O(10^5)$. *Such computations can, however, be made computationally effective by using a special type of polynomials, called* ***orthogonal polynomials****.*

Definition 10.5. The set of functions $\{\phi_0, \phi_1, \ldots, \phi_n\}$ in $[a, b]$ is called a set of **orthogonal functions**, with respect to a weight function $w(x)$, if

$$\int_a^b w(x)\phi_j(x)\phi_i(x)\,dx = \begin{cases} 0 & \text{if } i \neq j \\ C_j & \text{if } i = j \end{cases}$$

where C_j is a real positive number.

Furthermore, if $C_j = 1$, $j = 0, 1, \ldots, n$, then the orthogonal set is called an **orthonormal set**.

Using this interesting property, least-squares computations can be more numerically effective, as shown below. Without any loss of generality, **let's assume that** $w(x) = 1$.

Idea: The idea is to find a least-squares approximation of $f(x)$ on $[a, b]$ by means of a polynomial of the form

$$P_n(x) = a_0\phi_0(x) + a_1\phi_1(x) + \cdots + a_n\phi_n(x),$$

where $\{\phi_n\}_{k=0}^{n}$ is a set of orthogonal polynomials. That is, the basis for generating $P_n(x)$ in this case is a set of orthogonal polynomials.

The question now arises: Is it possible to write $P_n(x)$ as a linear combination of the orthogonal polynomials? The answer is ***yes*** and provided in the following. It can be shown [**Exercise**]:

Given the set of orthogonal polynomials $\{Q_i(x)\}_{i=0}^{n}$, a polynomial $P_n(x)$ of degree $\leq n$, can be written as:

$$P_n(x) = a_0Q_0(x) + a_1Q_1(x) + \cdots + a_nQ_n(x)$$

for *some* $a_0, a_1, \ldots, a_n$.

Finding the least-squares approximation of $f(x)$ on $[a, b]$ using orthogonal polynomials, then can be stated as follows:

Least-squares Approximation of a Function Using Orthogonal Polynomials

Given $f(x)$, continuous on $[a, b]$, find $a_0, a_1, \ldots, a_n$ using a polynomial of the form:

$$P_n(x) = a_0\phi_0(x) + a_1\phi_1(x) + \cdots + a_n\phi_n(x),$$

where

$$\{\phi_k(x)\}_{k=0}^{n}$$

is a given set of orthogonal polynomials on $[a, b]$, such that the error function:

$$E(a_0, a_1, \ldots, a_n) = \int_a^b [f(x) - (a_0\phi_0(x) + \cdots a_n\phi_n(x))]^2 \, dx$$

is minimized.

As before, we set

$$\frac{\partial E}{\partial a_i} = 0, \quad i = 0, 1, \ldots, n.$$

Now

$$\frac{\partial E}{\partial a_0} = -2\int_a^b \phi_0(x)[f(x) - a_0\phi_0(x) - a_1\phi_1(x) - \cdots - a_n\phi_n(x)] \, dx.$$

Setting this equal to zero, we get

$$\int_a^b \phi_0(x) f(x) \, dx = \int_a^b (a_0\phi_0(x) + \cdots + a_n\phi_n(x))\phi_0(x) \, dx.$$

Since, $\{\phi_k(x)\}_{k=0}^{n}$ is an orthogonal set, we have,

$$\int_a^b \phi_0^2(x) \, dx = C_0,$$

and

$$\int_a^b \phi_0(x)\phi_{i(x)} \, dx = 0, \quad i \neq 0.$$

Applying the above orthogonal property, we see from above that

$$\int_a^b \phi_0(x) f(x) \, dx = C_0 a_0.$$

That is,

$$a_0 = \frac{1}{C_0}\int_a^b \phi_0(x) f(x) \, dx.$$

Similarly,

$$\frac{\partial E}{\partial a_1} = -2\int_a^b \phi_1(x)[f(x) - a_0\phi_0(x) - a_1\phi_1(x) - \cdots - a_n\phi_n(x)] \, dx.$$

Again from the orthogonal property of $\{\phi_j(x)\}_{j=0}^n$ we have

$$\int_a^b \phi_1^2(x)\,dx = C_1 \text{ and } \int_a^b \phi_1(x)\phi_i(x)\,dx = 0, \quad i \neq 1,$$

so, setting $\frac{\partial E}{\partial a_1} = 0$, we get

$$a_1 = \frac{1}{C_1}\int_a^b \phi_1(x) f(x)\,dx$$

In general, we have

$$a_k = \frac{1}{C_k}\int_a^b \phi_k(x) f(x)\,dx, \quad k = 0, 1, \ldots, n,$$

where

$$C_k = \int_a^b \phi_k^2(x)\,dx.$$

10.1.4 Expressions for a_k with Weight Function $w(x)$.

If the weight function $w(x)$ is included, then a_k is modified to

$$a_k = \frac{1}{C_k}\int_a^b w(x) f(x)\phi_k(x)\,dx, \quad k = 0, 1, \ldots, n$$

Where $C_k = \int_a^b = \int_a^b w(x)\phi_k^2(x)dx$

Algorithm 10.6 (Least-Squares Approximation Using Orthogonal Polynomials).

Inputs: $f(x)$ - A continuous function $f(x)$ on $[a, b]$
$w(x)$ - A weight function (an integrable function on $[a, b]$).
$\{\phi_k(x)\}_{k=0}^n$ - A set of n orthogonal functions on $[a, b]$.

Output: The coefficients $a_0, a_1, \ldots, a_n$ such that

$$\int_a^b w(x)[f(x) - a_0\phi_0(x) - a_1\phi_1(x) - \cdots - a_n\phi_n(x)]^2\,dx$$

is minimized.

Step 1. Compute C_k, $k = 0, 1, \ldots, n$ as follows:
For $k = 0, 1, 2, \ldots, n$ do
$C_k = \int_a^b w(x)\phi_k^2(x)\,dx$
End

Step 2. Compute a_k, $k = 0, \ldots, n$ as follows:
For $k = 0, 1, 2, \ldots, n$ do
$a_k = \frac{1}{C_k} \int_a^b w(x) f(x) \phi_k(x)\, dx$
End

Generating a Set of Orthogonal Polynomials

The next question is: How to generate a set of orthogonal polynomials on $[a, b]$, with respect to a weight function $w(x)$?

The classical **Gram-Schmidt process** (see **Chapter 7**), can be used for this purpose. Starting with the set of monomials $\{1, x, \ldots, x^n\}$, we can generate the orthonormal set of polynomials $\{Q_k(x)\}$.

10.2 Least-Squares Approximation Using Legendre's Polynomials

Recall from the last chapter that the first few Legendre polynomials are given by:

$$\begin{cases} \phi_0(x) & = 1 \\ \phi_1(x) & = x \\ \phi_2(x) & = x^2 - \frac{1}{3} \\ \phi_3(x) & = x^3 - \frac{3}{5}x \\ \text{etc.} \end{cases}$$

To use the Legendre polynomials to approximate least-squares solution of a function $f(x)$, we set $w(x) = 1$, and $[a, b] = [-1, 1]$ and compute

- C_k, $k = 0, 1, \ldots, n$ by **Step 1** of Algorithm 10.6:
$$C_k = \int_{-1}^{1} \phi_k^2(x) dx$$
and
- a_k, $k = 0, 1, \ldots, n$ by **Step 2** of Algorithm 10.6:
$$a_k = \frac{1}{C_k} \int_{-1}^{1} f(x) \phi_k(x) dx.$$

The least-squares polynomial will then be given by $P_n(x) = a_0Q_0(x) + a_1Q_1(x) + \cdots + a_nQ_n(x)$.

A few C_k's are now listed below:

$$\begin{cases} C_0 &= \int_{-1}^{1} \phi_0^2(x)\,dx = \int_{-1}^{1} 1\,dx = 2 \\ C_1 &= \int_{-1}^{1} \phi_1^2(x)\,dx = \int_{-1}^{1} x^2\,dx = \frac{2}{3} \\ C_2 &= \int_{-1}^{1} \phi_2^2(x)\,dx = \int_{-1}^{1} \left(x^2 - \frac{1}{3}\right)^2\,dx = \frac{8}{45}. \end{cases}$$

and so on.

Example 10.7

Find linear and quadratic least-squares approximation to $f(x) = e^x$ using **Legendre polynomials**.

Linear Approximation: $P_1(x) = a_0\phi_0(x) + a_1\phi_1(x)$

$$\phi_0(x) = 1, \quad \phi_1(x) = x$$

Step 1. Compute $\mathbf{C_0}$ and $\mathbf{C_1}$:

$$C_0 = \int_{-1}^{1} \phi_0^2(x)\,dx = \int_{-1}^{1} dx = [x]_{-1}^{1} = 2$$

$$C_1 = \int_{-1}^{1} \phi_1^2(x)\,dx = \int_{-1}^{1} x^2\,dx = \left[\frac{x^3}{3}\right]_{-1}^{1} = \frac{1}{3} + \frac{1}{3} = \frac{2}{3}.$$

Step 2. Compute a_0 and a_1:

$$\begin{aligned} a_0 &= \frac{1}{C_0}\int_{-1}^{1} \phi_0(x)e^x dx \frac{1}{2}\int_{-1}^{1} e^x dx = \frac{1}{2}\left(e - \frac{1}{e}\right) \\ a_1 &= \frac{1}{C_1}\int_{-1}^{1} f(x)\phi_1(x)dx \\ &= \frac{3}{2}\int_{-1}^{1} e^x x dx = \frac{3}{e}. \end{aligned}$$

The **linear least-squares polynomial**:

$$\begin{aligned} P_1(x) &= a_0\phi_0(x) + a_1\phi_1(x) \\ &= \frac{1}{2}\left[e - \frac{1}{e}\right] + \frac{3}{e}x \end{aligned}$$

Accuracy Check:

$$P_1(0.5) = \frac{1}{2}\left[e - \frac{1}{e}\right] + \frac{3}{e}\cdot 0.5 = 1.7270$$
$$e^{0.5} = 1.6487$$

Relative error: $\frac{|1.7270-1.6487|}{|1.6487|} = 0.0475.$

Quadratic Approximation: $P_2(x) = a_0\,\phi_0(x) + a_1\,\phi_1(x) + a_2\,\phi_2(x)$

$$a_0 = \frac{1}{2}\left(e - \frac{1}{e}\right), \quad a_1 = \frac{3}{e} \text{ (\textit{Already computed above})}$$

Step 1. Compute

$$\begin{aligned} C_2 &= \int_{-1}^{1} \phi_2^2(x)\,dx = \int_{-1}^{1} \left(x^2 - \frac{1}{3}\right)^2 dx \\ &= \left(\frac{x^5}{5} - \frac{2}{3}\cdot\frac{x^3}{3} + \frac{1}{a}\,x\right)_{-1}^{1} = \frac{8}{45} \end{aligned}$$

Step 2. Compute

$$\begin{aligned} a_2 &= \frac{1}{C_2}\int_{-1}^{1} e^x \phi_2(x)\,dx \\ a_2 &= \frac{45}{8}\int_{-1}^{1} e^x\left(x^2 - \frac{1}{3}\right) dx \\ &= e - \frac{7}{e} \end{aligned}$$

The quadratic least-squares polynomial:

$$P_2(x) = \frac{1}{2}\left(e - \frac{1}{e}\right) + \frac{3}{e}\,x + \left(e - \frac{7}{e}\right)\left(x^2 - \frac{1}{3}\right)$$

Accuracy Check:

$$P_2(0.5) = 1.7151$$

$$e^{0.5} = 1.6487$$

Relative error: $\frac{|1.7151-1.6487|}{|1.6487|} = 0.0403.$

Compare this relative error with that obtained earlier with an non-orthogonal polynomial of degree 2.

10.3 Chebyshev polynomials: Another wonderful family of orthogonal polynomials

Definition 10.8. The set of polynomials defined by

$$T_n(x) = \cos[n \arccos x], \quad n \geq 0$$

on $[-1, 1]$ are called the **Chebyshev polynomials**.

To see that $T_n(x)$ is a polynomial of degree n in our familiar form, we derive a **recursive relation** by noting that

$T_0(x) = 1$ (the **Chebyshev polynomial** of degree zero).

$T_1(x) = x$ (the **Chebyshev polynomial** of degree 1).

10.3.1 A Recursive Relation for Generating Chebyshev Polynomials

Substitute $\theta = \arccos x$. Then,

$$\begin{aligned}
T_n(x) &= \cos(n\theta), \quad 0 \leq \theta \leq \pi. \\
T_{n+1}(x) &= \cos(n+1)\theta = \cos n\theta \cos\theta - \sin n\theta \sin\theta \\
T_{n-1}(x) &= \cos(n-1)\theta = \cos n\theta \cos\theta + \sin n\theta \sin\theta
\end{aligned}$$

Adding the last two equations, we obtain

$$T_{n+1}(x) + T_{n-1}(x) = 2\cos n\theta \cos\theta$$

The right-hand side still does not look like a polynomial in x. But note that $\cos\theta = x$

So,

$$\begin{aligned}
T_{n+1}(x) &= 2\cos n\theta \cos\theta - T_{n-1}(x) \\
&= 2xT_n(x) - T_{n-1}(x).
\end{aligned}$$

This above is a three-term recurrence relation to generate the Chebyshev Polynomials.

Three-Term Recurrence Formula for Chebyshev Polynomials

$$T_0(x) = 1, \; T_1(x) = x$$

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x), \quad n \geq 1.$$

Using this recursive relation, the Chebyshev polynomials of the successive degrees can be generated:

$$\begin{aligned} n = 1: &\quad T_2(x) = 2xT_1(x) - T_0(x) = 2x^2 - 1, \\ n = 2: &\quad T_3(x) = 2xT_2(x) - T_1(x) = 2x(2x^2 - 1) - x = 4x^3 - 3x \end{aligned}$$

and so on.

10.3.2 The orthogonal property of the Chebyshev polynomials

We now show that *Chebyshev polynomials are orthogonal with respect to the weight function*

$$w(x) = \frac{1}{\sqrt{1-x^2}} \text{ in the interval } [-1, 1].$$

To demonstrate the orthogonal property of these polynomials, we show that

Orthogonal Property of the Chebyshev Polynomials

$$\int_1^{-1} \frac{T_m(x)T_n(x)}{\sqrt{1-x^2}}\,dx = \begin{cases} 0 & \text{if } m \neq n \\ \frac{\pi}{2} & \text{if } m = n. \end{cases}$$

First,

$$\begin{aligned} &\int_{-1}^{1} \frac{T_m(x)T_n(x)}{\sqrt{1-x^2}}\,dx, \quad m \neq n. \\ &\quad = \int_{-1}^{1} \frac{\cos(m \arccos x)\cos(n \arccos x)}{\sqrt{1-x^2}}\,dx \end{aligned}$$

Since

$$\begin{aligned} \theta &= \arccos x \\ d\theta &= -\frac{1}{\sqrt{1-x^2}} dx, \end{aligned}$$

The above integral becomes:

$$-\int_{\pi}^{0} \cos m\theta \cos n\theta d\theta$$
$$= \int_{0}^{\pi} \cos m\theta \cos n\theta d\theta$$

Now, $\cos m\theta \cos n\theta$ can be written as $\frac{1}{2}[\cos(m+n)\theta + \cos(m-n)\theta]$

So,

$$\begin{aligned}
&\int_0^\pi \cos m\theta \cos n\theta \, d\theta \\
&= \frac{1}{2}\int_0^\pi \cos(m+n)\theta \, d\theta + \frac{1}{2}\int_0^\pi \cos(m-n)\theta \, d\theta \\
&= \frac{1}{2}\left[\frac{1}{(m+n)}\sin(m+n)\theta\right]_0^\pi + \frac{1}{2}\left[\frac{1}{(m-n)}\sin(m-n)\theta\right]_0^\pi \\
&= 0.
\end{aligned}$$

Similarly, it can be shown [**Exercise**] that

$$\int_1^{-1} \frac{T_n^2(x)\, dx}{\sqrt{1-x^2}} = \frac{\pi}{2} \text{ for } n \geq 1.$$

10.4 The Least-Square Approximation using Chebyshev Polynomials

As before, the Chebyshev polynomials can be used to find least-squares approximations to a function $f(x)$ as stated below.

In Algorithm 10.6, set $w(x) = \frac{1}{\sqrt{1-x^2}}$,

$[a, b] = [-1, 1]$,

and $\phi_k(x) = T_k(x)$,

Then, it is easy to see that using the orthogonal property of Chebyshev polynomials:

$$C_0 = \int_{-1}^{1} \frac{T_0^2(x)}{\sqrt{1-x^2}} dx = \pi$$
$$C_k = \int_{-1}^{1} \frac{T_k^2(x)}{\sqrt{1-x^2}} dx = \frac{\pi}{2}, \qquad k = 1, \ldots, n$$

Thus, from Step 2 of Algorithm 10.6, we have

$$a_0 = \frac{1}{\pi} \int_{-1}^{1} \frac{f(x)dx}{\sqrt{1-x^2}}$$
$$\text{and } a_i = \frac{2}{\pi} \int_{-1}^{1} \frac{f(x)dx}{\sqrt{1-x^2}}.$$

The least-squares approximating polynomial $P_n(x)$ of $f(x)$ using Chebyshev polynomials is given by:

$$P_n(x) = a_0 T_0(x) + a_1 T_1(x) + \cdots + a_n T_n(x)$$

where

$$a_i = \frac{2}{\pi} \int_{-1}^{1} \frac{f(x) T_i(x)\, dx}{\sqrt{1-x^2}}, \quad i = 1, \cdots, n$$

and

$$a_0 = \frac{1}{\pi} \int_{-1}^{1} \frac{f(x)\, dx}{\sqrt{1-x^2}}$$

Example 10.9

Find a **linear** least-squares approximation of $f(x) = e^x$ using Chebyshev polynomials.

Here

$$P_1(x) = a_0 \phi_0(x) + a_1 \phi_i(x) = a_0 T_0(x) + a_1 T_1(x) = a_0 + a_1 x,$$

where

$$a_0 = \frac{1}{\pi} \int_{-1}^{1} \frac{e^x\, dx}{\sqrt{1-x^2}} \approx 1.2660$$
$$a_1 = \frac{2}{\pi} \int_{-1}^{1} \frac{x e^x}{\sqrt{1-x^2}}\, dx \approx 1.1303$$

Thus, $P_1(x) = 1.2660 + 1.1303x$

Accuracy Check:

$$P_1(0.5) = 1.8312;$$
$$e^{0.5} = 1.6487$$

Relative error: $\frac{|1.6487-1.8312|}{1.6487} = 0.1106$.

10.5 Monic Chebyshev Polynomials

Note that $T_k(x)$ is a Chebyshev polynomial of degree k with the leading coefficient 2^{k-1}, $k \geq 1$. Thus we can generate a set of monic Chebyshev polynomials from the polynomials $T_k(x)$ as follows:

- The **Monic Chebyshev Polynomials**, $\tilde{T}_k(x)$, are then given by
$$\tilde{T}_0(x) = 1, \quad \tilde{T}_k(x) = \frac{1}{2^{k-1}} T_k(x), \quad k \geq 1.$$
- The k zeros of $\tilde{T}_k(x)$ are easily calculated [**Exercise**]:
$$\tilde{x}_j = \cos\left(\frac{2j-1}{2k}\pi\right), \quad j = 1, 2, \ldots, k.$$
- The maximum or minimum values of $\tilde{T}_k(x)$ [**Exercise**] occur at
$$\tilde{x}_j = \cos\left(\frac{j\pi}{k}\right),$$
and
$$\tilde{T}_k(\tilde{x}_j) = \frac{(-1)^j}{2^{k-1}}, \quad j = 0, 1, \ldots, k.$$

10.6 Minimax Polynomial Approximations with Chebyshev Polynomials

As seen above the Chebyshev polynomials can, of course, be used to find least-squares polynomial approximations. *However, these polynomials have several other wonderful polynomial approximation properties.* First, we state the **minimax property of the Chebyshev polynomial**s.

Minimax Property of the Chebyshev Polynomials

If $P_n(x)$ is any monic polynomial of degree n, then
$$\frac{1}{2^{n-1}} = \max_{x \in [-1,1]} |\tilde{T}_n(x)| \leq \max_{x \in [-1,1]} |P_n(x)|.$$
Moreover, this happens when
$$P_n(x) \equiv \tilde{T}_n(x).$$

Interpretation: The above results say that the absolute maximum of the monic Chebyshev polynomial of degree n, $\tilde{T}_n(x)$, over $[-1,1]$ is less than or equal to that of any polynomial $P_n(x)$ over the same interval.

Proof. By contradiction [**Exercise**].

10.6.1 Choosing the interpolating nodes with the Chebyshev Zeros

Recall from **Chapter 6** that error in **polynomial interpolation** of a function $f(x)$ with the nodes, $x_0, x_1, \ldots, x_n$ by a polynomial $P_n(x)$ of degree at most n is given by

$$E = f(x) - P(x) = \frac{f^{n+1}(\xi)}{(n+1)!}\ \Psi(x),$$

where $\Psi(x) = (x - x_0)(x - x_1)\cdots(x - x_n)$.

The question is: *How to choose these* $(n+1)$ *nodes* $x_0, x_1, \ldots, x_n$ *so that* $|\Psi(x)|$ *is minimized in* $[-1,1]$*?*

The answer can be given from the above minimax property of the monic Chebyshev polynomials.

Note that $\Psi(x)$ is a monic polynomial of degree $(n+1)$.

So, by the *minimax property of the Chebyshev polynomials*, we have

$$\max_{x\in[-1,1]} |\tilde{T}_{n+1}(x)| \leq \max_{x\in[-1,1]} |\Psi(x)|.$$

Thus,

• The maximum value of $\Psi(x) = (x - x_0)(x - x_1)\cdots(x - x_n)$ in $[-1,1]$ is smallest when $x_0, x_1, \ldots, x_n$ are chosen as the $(n+1)$ zeros of the $(n+1)$th degree monic Chebyshev polynomials $\tilde{T}_{n+1}(x)$.

That is, when $x_0, x_1, \ldots, x_n$ are chosen as:

$$\tilde{x}_{k+1} = \cos\frac{(2k+1)}{2(n+1)}\pi, k = -1, 0, 1, \ldots, n-1.$$

• The smallest absolute maximum value of $\Psi(x)$ with $x_0, x_1, \ldots, x_n$ as chosen above, is: $\frac{1}{2^n}$.

10.6.2 Working with an Arbitrary Interval

If the interval is $[a,b]$, different from $[-1,1]$, then, the zeros of $\tilde{T}_{n+1}(x)$ need to be shifted by using the transformation:

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

Example 10.10

Let the interpolating polynomial be of degree at most 2 and the interval be $[1.5, 2]$.

The three zeros of $\tilde{T}_3(x)$ in $[-1, 1]$ are given by

$$\tilde{x}_0 = \cos\frac{\pi}{6}, \quad \tilde{x}_1 = \cos\frac{\pi}{6}, \text{ and } \tilde{x}_2 = \cos\frac{pi}{2}.$$

These zeros are to be shifted using transformation:

$$x_{\text{new } i} = \frac{1}{2}\left[(2-1.5)\tilde{x}_i + (2+1.5)\right], \qquad i = 0, 1, 2.$$

10.6.3 Use of Chebyshev Polynomials to Economize Power Series

Power Series Economization

Let $P_n(x) = a_0 + a_1 x + \cdots + a_n x^n$ be a polynomial of degree n obtained by truncating a power series expansion of a continuous function on $[a, b]$. The problem is to find a polynomial $P_r(x)$ of degree r $(< n)$ such that

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

is as small as possible.

We first consider $r = n - 1$; that is, the problem of approximating $P_n(x)$ by a polynomial $P_{n-1}(x)$ of degree $n - 1$. If the total error (sum of truncation error and error of approximation) is still within an acceptable tolerance, then we consider approximation by a polynomial of $P_{n-1}(x)$ and the process is continued until the accumulated error exceeds the tolerance.

We will show below how minimax property of the Chebyshev polynomials can be gainfully used. First note that $|\frac{1}{a_n}P_n(x) - P_{n-1}(x)|$ is a monic polynomial. So, by the minimax property, we have

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

Thus, if we choose

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

Then the minimum value of the maximum error of approximating $P_n(x)$ by $P_{n-1}(x)$ over $[-1, 1]$ is given by:

$$\max_{-1 \leq x \leq 1} |P_n(x) - P_{n-1}(x)| = \frac{|a_n|}{2^{n-1}}.$$

If this quantity, $\frac{|a_n|}{2^{n-1}}$, plus error due to the truncation of the power series is within the permissible tolerance ϵ, we can then repeat the process by constructing $P_{n-2}(x)$ from $P_{n-1}(x)$ as above. The process can be continued until the accumulated error exceeds the error tolerance ϵ.

So, the process can be summarized as follows:

Power Series Economization Process by Chebyshev Polynomials

Inputs: (i) $f(x)$ - $(n+1)$ times differentiable function on $[a, b]$

(ii) $n =$ positive integer

Outputs: An economized power series of $f(x)$.

Step 1. Obtain $P_n(x) = a_0 + a_1 x^n + \cdots + a_n x^n$ by truncating the power series expansion of $f(x)$.

Step 2. Find the truncation error.

$$E_{TR}(x) = \text{ Reminder after } n \text{ terms } = \frac{|f^{(n+1)}(\xi(x))||x^{n+1}|}{(n+1)!}$$

and compute the upper bound of $|E_{TR}(x)|$ for $-1 \leq x \leq 1$.

Step 3. Compute $P_{n-1}(x)$:

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

Step 4. Check if the maximum value of the total error $\left(|E_{TR}| + \frac{|a_n|}{2^{n-1}}\right)$ is less than ϵ. If so, approximate $P_{n-1}(x)$ by $P_{n-2}(x)$. Compute $P_{n-2}(x)$ as:

$$P_{n-2}(x) = P_{n-1}(x) - a_{n-1} \tilde{T}_{n-1}(x)$$

Step 5. Compute the error of the current approximation:

$$|P_{n-1}(x) - P_{n-2}(x)| = \frac{|a_{n-1}|}{2^{n-2}}.$$

Step 6. Compute the accumulated error:

$$\left(|E_{TR}| + |\frac{a_n}{2^{n-1}}| + |\frac{a_{n-1}}{2^{n-2}}|\right)$$

If this error is still less than ϵ, continue until the accumulated error exceeds ϵ.

Example 10.11

Find the economized power series for $\cos x = 1 - \frac{1}{2}x^2 + \frac{x^4}{4!} - \cdots$ for $0 \leq x \leq 1$ with a tolerance $\epsilon = 0.05$.

Step 1. Let's first try with $n = 4$.

$$P_4(x) = 1 - \frac{1}{2}x^2 + 0 \cdot x^3 + \frac{x^4}{4!}$$

Step 2. Truncation Error.

$$T_R(x) = \frac{f^5(\xi(x))}{5!}x^5$$

since $f^{(5)}(x) = -\sin x$ and $|\sin x| \leq 1$ for all x, we have

$$|T_R(x)| \leq \frac{1}{120}(1)(1)^5 = 0.0083.$$

Step 3. Compute $P_3(x)$:

$$\begin{aligned} P_3(x) &= P_4(x) - a_4\tilde{T}_4(x) \\ &= 1 - \frac{1}{2}x^2 + \frac{x^4}{4!} - \frac{1}{4!}\left(x^4 - x^2 + \frac{1}{8}\right) \\ &= -0.4583x^2 + 0.9948 \end{aligned}$$

Step 4. Maximum absolute accumulated error = Maximum absolute truncation error + maximum absolute Chebyshev error $= 0.0083 + \frac{1}{24} \cdot \frac{1}{8} = 0.0135$ which is greater than $\epsilon = 0.05$.

Thus, the **economized power series** of $f(x) = \cos x$ for $0 \leq x \leq 1$ with an error tolerance of $\epsilon = 0.05$ is the 3^{rd} degree polynomial:

$$P_3(x) = -0.4583x^2 + 0.9948.$$

Accuracy Check: We will check below how accurate this approximation is for three values of x in $0 \leq x \leq 1$: two extreme values $x = 0$ and $x = 1$, and one intermediate value $x = 0.5$.

For x=0.5.

$$\begin{aligned} \cos(0.5) &= 0.8776 \\ P_3(0.5) &= 0.8802 \\ \text{Error} &= |\cos(0.5) - P_3(0.5)| = |-0.0026| = 0.0026. \end{aligned}$$

For x=0.

$$\cos(0) = 1; \quad P_3(0) = 0.9948$$

$$\text{Error} = |\cos(0) - P_3(0)| = 0.0052.$$

For x=1.

$$\cos(1) = 0.5403; \quad P_3(1) = 0.5365$$
$$\text{Error} = |\cos(1) - P_3(1)| = 0.0038.$$

Remark: These errors are all much less than 0.05 even though theoretical error-bound, 0.0135, exceeded it.