

Discrete Least-squares Approximations

Given a set of data points $(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)$, 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 fit best 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 data points $(x_k, y_i), i = 1, \dots, m$, find an algebraic polynomial $P_n(x) = a_0 + a_1x + \dots + a_nx^n$ ($n < m$) such that the error in the least-squares sense is minimized; that is, $E = \sum_{i=1}^m (y_i - a_0 - a_1x_i - \dots - a_nx_i^n)^2$ is minimum.

For E to be minimum, we must have

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

Now,

$$\begin{aligned} \frac{\partial E}{\partial a_0} &= -2 \sum_{i=1}^m (y_i - a_0 - a_1x_i - \dots - a_nx_i^n) \\ \frac{\partial E}{\partial a_1} &= -2 \sum_{i=1}^m x_i (y_i - a_0 - a_1x_i - \dots - a_nx_i^n) \\ &\vdots \\ \frac{\partial E}{\partial a_n} &= -2 \sum_{i=1}^m x_i^n (y_i - a_0 - a_1x_i - \dots - a_nx_i^n) \end{aligned}$$

Setting these equations to be zero we have

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

Set now $\sum_{i=1}^m x_i^k = s_k$, $k = 0, 1, \dots, 2n$, and denoting the right hand side entries as b_0, \dots, b_n , the above equation can be written as:

$$\begin{aligned} s_0 a_0 + s_1 a_1 + \dots + s_n a_n &= b_0 \quad (\text{Note that } \sum_{i=1}^m x_i^0 = s_0 = m). \\ s_1 a_0 + s_2 a_1 + \dots + s_{n+1} a_n &= b_1 \\ &\vdots \\ s_n a_0 + s_{n+1} a_1 + \dots + s_{2n} a_n &= b_n \end{aligned}$$

This is a system of $(n + 1)$ equations in $(n + 1)$ unknowns a_0, a_1, \dots, 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:

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

or

$$sa = b$$

where

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

Define

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

Then the above system has the form:

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

The matrix V is known as the **Vandermonde matrix**, and it can be shown [**Exercise**] that it 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 0.1 (*Existence and uniqueness of Discrete Least-Squares Solutions*).
 Let $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ be n distinct points. Then the discrete least-square approximation problem has a unique solution.

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.

The problem can be stated as follows:

Least-Square Approximations of a Function Using Standard Polynomials

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

$$P_n(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$$

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, \dots, a_n , we denote this by $E(a_0, a_1, \dots, a_n)$.

For minimization, we must have

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

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

Setting up the Normal Equations

Since

$$E = \int_a^b [f(x) - (a_0 + a_1x + a_2x^2 + \cdots + a_nx^n)]^2 dx$$

We have

$$\frac{\partial E}{\partial a_0} = -2 \int_a^b [f(x) - a_0 - a_1x - a_2x^2 - \cdots - a_nx^n] dx$$

$$\frac{\partial E}{\partial a_1} = -2 \int_a^b x[f(x) - a_0 - a_1x - a_2x^2 - \cdots - a_nx^n] dx$$

⋮

$$\frac{\partial E}{\partial a_n} = -2 \int_a^b x^n[f(x) - a_0 - a_1x - a_2x^2 - \cdots - a_nx^n] dx$$

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

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

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

$$i = 0 : a_0 \int_a^b 1 dx + a_1 \int_a^b x dx + a_2 \int_a^b x^2 dx + \cdots + a_n \int_a^b x^n dx = \int_a^b f(x) dx$$

$$i = 1 : a_0 \int_a^b x dx + a_1 \int_a^b x^2 dx + a_2 \int_a^b x^3 dx + \cdots + a_n \int_a^b x^n dx = \int_a^b x f(x) dx$$

⋮

$$i = n : a_0 \int_a^b x^n dx + a_1 \int_a^b x^{n+1} dx + a_2 \int_a^b x^{n+2} dx + \cdots + a_n \int_a^b x^{2n} dx = \int_a^b x^n f(x) dx$$

Denote

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

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

$$a_0 s_0 + a_1 s_1 + a_2 s_2 + \cdots + a_n s_n = b_0$$

$$a_0 s_1 + a_1 s_2 + a_2 s_3 + \cdots + a_{n+1} s_{n+1} = b_1$$

⋮

$$a_0 s_n + a_1 s_{n+1} + a_2 s_{n+2} + \cdots + a_{2n} s_{2n} = b_n.$$

or in matrix notation

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

Denote

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

Then we have the linear system:

$$Sa = b$$

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

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

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

Thus, in this case the matrix of the normal equations

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

which is a **Hilbert Matrix**. It is well-known to be **ill-Conditioned**.

Algorithm: **Least-Squares Approximation using Polynomials**

Inputs:

- (i) $f(x)$ - A continuous function on $[a, b]$.
- (ii) n - The degree of the desired least-square polynomial

Output: The coefficients a_0, a_1, \dots, a_n of the desired least-squares polynomial: $P_n(x) = a_0 + a_1x + \dots + a_nx^n$.

Step 1. Compute s_0, s_1, \dots, s_{2n}

For $i = 0, 1, \dots, 2n$ do

$$s_i = \int_a^b x^i dx$$

End

Step 2 Compute b_0, b_1, \dots, b_n :

For $i = 0, 1, \dots, n$ do

$$b_i = \int_a^b x^i f(x) dx$$

End.

Step 3 Form the matrix S and the vector b .

$$S = \begin{pmatrix} s_0 & s_1 & \dots & s_n \\ s_1 & s_2 & \dots & s_{n+1} \\ \vdots & \vdots & & \vdots \\ s_n & s_{n+1} & \dots & s_{2n} \end{pmatrix}$$

$$b = \begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_n \end{pmatrix}.$$

Step 4 Solve the $(n+1) \times (n+1)$ system of equations:

$$S a = b,$$

$$\text{where } a = \begin{pmatrix} a_0 \\ \vdots \\ a_n \end{pmatrix}.$$

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

Linear Approximation:

$$n = 1; P_1(x) = a_0 + a_1x$$

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

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

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

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

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

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

The normal system of equations 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 $P_1(x) = 1.1752 + 1.1037x$

Check Accuracy:

$$P_1(0.5) = 1.7270$$

$$e^{0.5} = 1.6487$$

$$\text{Relative Error} = \frac{|1.6487 - 1.7270|}{|1.7270|} = 0.0453.$$

Quadratic Fitting:

$$n = 2$$

$$P_2(x) = a_0 + a_1x + a_2x^2$$

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

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

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

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

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

$$b_2 = \int_{-1}^1 x^2 e^x dx = e - \frac{5}{e} \approx 0.8789.$$

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 is: $a_0 = 0.9963$, $a_1 = 1.1037$, $a_2 = 0.5368$.

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

Check the accuracy:

$$P_2(0.5) = 1.6889$$

$$e^{0.5} = 1.6487$$

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

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

Linear Fit:

$$P_1(x) = a_0 + a_1x$$

$$s_0 = \int_0^1 dx = 1$$

$$s_1 = \int_0^1 x dx = \frac{1}{2}$$

$$s_2 = \int_0^1 x^2 dx = \frac{1}{3}$$

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

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

The **normal equations** are:

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

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

Check Accuracy:

$$f(0.5) = 8.75; P_1(0.5) = 8.833$$

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

Quadratic Least-Square Approximation: $P_2(x) = a_0 + a_1x + a_2x^2$

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

$$b_0 = \frac{53}{6}, \quad b_1 = \frac{59}{12}.$$

$$b_2 = \int_0^1 x^2(x^2 + 5x + 6)dx = \int_0^1 (x^4 + 5x^3 + 6x^2)dx = \frac{1}{5} + \frac{5}{4} + \frac{6}{3} = \frac{69}{20}$$

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

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

Use of Orthogonal Polynomials in Least-squares Approximations

The least-squares approximation using 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. The set of functions $\phi_0, \phi_1, \dots, \phi_n$ 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 $C_j = 1, j = 0, 1, \dots, n$, 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 an approximation of $f(x)$ on $[a, b]$ by means of a polynomial of the form

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

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.

Least-squares Approximation of a Function Using Orthogonal Polynomials

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

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

where

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

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

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

is minimized.

As before, we set

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

Now

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

Setting this equal to zero, we get

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

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

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

and

$$\int_a^b \phi_0(x)\phi_i(x) dx = 0, i \neq 0.$$

Applying the above orthogonal property, we see from above that

$$\int_a^b \phi_0(x)f(x)dx = C_0a_0.$$

That is,

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

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

The orthogonal property of $\{\phi_j(x)\}_{j=0}^n$ implies that

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

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,

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

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

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

If the weight function $w(x)$ is included, we obtain

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

Algorithm: Least-Squares Approximation Using Orthogonal Polynomials

Inputs:

$f(x)$ - A continuous function 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, \dots, a_n such that

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

is minimized.

Step 1. Compute $C_k, k = 0, 1, \dots, n$ as follows:

For $k = 0, 1, 2, \dots, n$ do

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

End

Step 2. Compute $a_k, k = 0, \dots, n$ as follows:

For $k = 0, 1, 2, \dots, n$ do

$$a_k = \frac{1}{C_k} \int_a^b w(x)f(x)\phi_k(x) dx$$

End

Least-Squares Approximation Using Legendre's Polynomials

Recall that the **Legendre Polynomials** $\{\phi_k(x)\}$ are given by

$$\phi_0(x) = 1$$

$$\phi_1(x) = x$$

$$\phi_2(x) = x^2 - \frac{1}{3}$$

$$\phi_3(x) = x^3 - \frac{3}{5}x$$

etc.

are orthogonal polynomials on $[-1, 1]$, with respect to the weight function $w(x) = 1$.

If these polynomials are used for least-squares approximation, then it is easy to see that

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

and so on.

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

Linear Approximation:

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

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

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

$$a_0 = \frac{1}{C_0} \int_{-1}^1 \phi_1(x) e^x dx$$

$$\text{So, } a_0 = \frac{1}{2} \int_{-1}^1 e^x dx = \frac{1}{2} [e^x]_{-1}^1 = \frac{1}{2} \left(e - \frac{1}{e} \right)$$

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

$$a_1 = \frac{3}{2} \int_{-1}^1 x e^x dx = \frac{3}{2} \left[\frac{2}{e} \right] = \frac{3}{e}$$

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:

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

$$e^{0.5} = 1.6487$$

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

Quadratic Approximation:

$$P_2(x) = a_0 \phi_0(x) + a_1 \phi_1(x) + a_2 \phi_2(x)$$

$$a_0 = \frac{1}{2} \left(e - \frac{1}{e} \right), \quad a_1 = \frac{3}{e}$$

$$\begin{aligned} C_2 &= \int_{-1}^1 \phi_2^2(x) dx = \int_{-1}^1 \left(x^2 - \frac{1}{3} \right)^2 dx \\ &= \left(\frac{x^5}{5} - \frac{2}{3} \cdot \frac{x^3}{3} + \frac{1}{9} x \right) \Big|_{-1}^1 = \frac{8}{45} \end{aligned}$$

$$a_2 = \frac{1}{C_2} \int_{-1}^1 e^x \phi_2(x) dx$$

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

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:

$$P_n(0.5) = 1.5868$$

$$e^{0.5} = 1.6487$$

Relative error

$$\frac{|1.5868 - 1.6487|}{|1.6487|} = 0.0375$$

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

Chebyshev polynomials: Another wonderful family of orthogonal polynomials

Definition: The set of polynomials defined by

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

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

$$T_0(x) = 1 \text{ (A polynomial of degree zero).}$$

$$T_1(x) = x \text{ (A polynomial of degree 1).}$$

A Recursive Relation for Generating Chebyshev Polynomials:

Substitute $\theta = \arccos x$

Then, $T_n(x) = \cos(n\theta)$, $0 \leq \theta \leq \pi$.

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

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

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,

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

or

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

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

$$n = 1 : T_2(x) = 2xT_1(x) - T_0(x) = 2x^2 - 1.$$

$$n = 2 : T_3(x) = 2xT_2(x) - T_1(x) = 2x(2x^2 - 1) - x = 4x^3 - 3x.$$

and so on.

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, consider

$$\begin{aligned} & \int_{-1}^1 \frac{T_m(x)T_n(x)dx}{\sqrt{1-x^2}}, \quad m \neq n. \\ &= \int_{-1}^1 \frac{\cos(\arccos x) \cos(n \arccos x)}{\sqrt{1-x^2}} dx \\ &= \int_0^\pi \cos m\theta \cos n\theta d\theta \quad (\text{By changing the variable from } x \text{ to } \theta \text{ with substitution of } \arccos x = \theta). \\ &= \frac{-1}{2} \int_0^\pi \cos(m+n)\theta d\theta + \frac{-1}{2} \int_0^\pi \cos(m-n)\theta d\theta \\ &= \frac{-1}{2} \left[\frac{1}{(m+n)} \sin(m+n)\theta \right]_0^\pi + \frac{-1}{2} \left[\frac{1}{(m-n)} \sin(m-n)\theta \right]_0^\pi \\ &= 0. \end{aligned}$$

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

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

Summarizing:

Orthogonal Property of the Chebyshev Polynomials

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

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.

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

$$P_n(x) = C_0T_0(x) + C_1T_1(x) + \cdots + C_nT_n$$

where

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

and

$$C_0 = \frac{1}{\pi} \int_{-1}^1 \frac{f(x)dx}{\sqrt{1-x^2}}$$

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

Here

$$P_1(x) = a_0\phi_0(x) + a_1\phi_1(x) = a_0T_0(x) + a_1T_1(x) = a_0 + a_1x,$$

where

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

$$a_1 = \frac{2}{\pi} \int_{-1}^1 \frac{xe^x}{\sqrt{1-x^2}} dx \approx 1.1303$$

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

Check the accuracy:

$$P_1(0.5) = 1.9175; e^{0.5} = 1.6487$$

Relative error $\frac{|1.6487 - 1.9175|}{1.9175} = 0.1402$

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

$$\tilde{T}_0(x) = 1, \tilde{T}_k(x) = \frac{1}{2^{k-1}}T_k(x), \quad k \geq 1.$$

- The k zeros of $\tilde{T}_k(x)$ are easily calculated [**Exercise**]:

$$x_j = \cos\left(\frac{2j-1}{2k}\pi\right), \quad j = 1, 2, \dots, k.$$

- The maximum or minimum values of $\tilde{T}_k(x)$ occur at $\tilde{x}_j = \cos\left(\frac{j\pi}{k}\right)$, and

$$\tilde{T}_k(\tilde{x}_j) = \frac{(-1)^j}{2^{k-1}}, \quad j = 0, 1, \dots, k.$$

Polynomial Approximations with Chebyshev the 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.** Some of them are stated below.

- The maximum absolute value of any monic polynomial of degree n over $[-1, 1]$ is always greater than or equal to that of $T_n(x)$ over the same interval; which is, by the last property, $\frac{1}{2^{n-1}}$.

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

$$P_n(x) = \tilde{T}_n(x).$$

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

- **Choosing the interpolating nodes with the Chebyshev Zeros.**

Recall that error in **polynomial interpolation** 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 $\Psi(x) = (x - x_0)(x - x_1) \dots (x - x_n)$.

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

The answer can be given from the last-mentioned property of the monic Chebyshev polynomials.

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

So, by the minimax property

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

That is, the maximum value of $\psi(x)$ is smallest when x_0, x_1, \dots, x_n are chosen as the $(n+1)$ zeros of $\tilde{T}_{n+1}(x)$ and this maximum value is $\frac{1}{2^n}$.

Choosing the Nodes for Minimizing Polynomial Interpolation error

To minimize the polynomial interpolation error, choose the nodes x_0, x_1, \dots, x_n as the $(n+1)$ zeros of the $(n+1)$ th degree monic Chebyshev polynomial.

Note. (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 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

$$\tilde{x}_1 = \cos \frac{\pi}{6}, \quad \tilde{x}_2 = \cos \frac{\pi}{2}, \quad \text{and} \quad \tilde{x}_3 = \cos \frac{5}{6}\pi.$$

These zeros are to be shifted using transformation:

$$x_{new} = \frac{1}{2}[(2-1.5)\tilde{x}_i + (2+1.5)]$$

- **Use of Chebyshev Polynomials to Economize Power Series**

Power Series Economization

Let $P_n(x) = a_0 + a_1x + \dots + a_nx^n$ 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

$$|P_n(x) - P_r(x)| < \epsilon,$$

where ϵ is a tolerance supplied by users.

The problem is easily solved by using the Minimax Property of the Chebyshev polynomials. 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

$$\max |P_n(x) - P_{n-1}(x)| \geq \frac{1}{a_n} \max |\tilde{T}_n(x)| = \frac{|a_n|}{2^{n-1}}$$

Thus, if we choose

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

then the minimum value of $\max |P_n(x) - P_{n-1}(x)| = \frac{|a_n|}{2^{n-1}}$.

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 tolerance ϵ .

So, the process can be summarized as follows:

Power Series Economization Process by Chebyshev Polynomials

- Obtain $P_n(x) = a_0 + a_1x^n + \dots + a_nx^n$ by truncating the power series expansion of $f(x)$.
- Find the error of truncation E_T^P
- Compute $P_{n-1}(x)$:

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

- Check if the total error $\left(|E_T^P| + \frac{|a_n|}{2^{n-1}}\right)$ is less than ϵ . If so, continue the process by decreasing the degree of the polynomials successively until the accumulated error becomes greater than ϵ .