Approximation Theory

Most functions cannot be evaluated exactly, even though we usually handle them as if they were completely known quantities. The simplest and most important of these are $\sqrt{x}$, e^x, $\ln x$, the trigonometric functions, and y^x; and there are many other functions that occur commonly in physics, engineering, and other areas.

In evaluating functions, by hand or using a computer, we are essentially limited to the elementary arithmetic operations $+$, $-$, $\times$, and $\div$. Combining these means we can evaluate polynomials and rational functions. All other functions must be evaluated using approximations based on polynomials or rational functions.

The Taylor polynomial is a relatively easy way to approximate a function to any desired level of accuracy; and often it is the only direct method of approximating the function. Nonetheless, a Taylor polynomial for $f(x)$ is usually a very inefficient approximation; if the approximation is to be used many times, then it should be replaced by a formula requiring less evaluation time. If a Taylor polynomial of some degree is being used, then there is usually another polynomial of much lower degree that will be of equal accuracy; and its lower degree will decrease both the evaluation time and the number of rounding errors in the evaluation process. The Taylor polynomial gives us an accurate, but inefficient way to approximate a function. Often it will be the only way to initially evaluate a function; then this approximation is used to find a more efficient approximation.

♣ **The Minimax Approximation Problem**. Let $f(x)$ be continuous on a given interval $[a, b]$. If $p(x)$ is a polynomial, then we are interested in measuring

$$E(f, p) = \max_{a \leq x \leq b} \{|f(x) - p(x)|\}$$

the worst possible error in the approximation of $f(x)$ by $p(x)$. For each degree $n \geq 0$, define

$$\rho_n(f) = \min_{a \leq x \leq b} E(f, p) = \min_{deg(p) \leq n} \max_{a \leq x \leq b} \{|f(x) - p(x)|\}$$

This is the smallest possible value for $E(p)$ that can be attained with a polynomial of degree $\leq n$. It is called the *minimax error*. The polynomial that gives this value is called the *minimax Polynomial Approximation* of order n, and it will be denoted by $m_n(x)$.

♣ **Discrete Least-Square Approximation**. Consider the problem of estimating the values of the function $f(x) = 2^x$ with a polynomial $p_1(x) = ax + b$ at the following points:

x_i :	0	1	2	3
y_i :	1	2	4	8

The problem of finding the equation of the best linear approximation in the absolute sense requires that values of a and b be found to minimize

$$\max_{i=1,2,3,4} \{|y_i - (ax_i + b)|\}$$

The *Least-square* approach to this problem involves determining the best approximating line $y = ax + b$, when the error involved is the sum of the squares of the differences between the values on the approximating line and given values. Therefore we need to find constants a and b that minimize

$$E(a,b) = \sum_{i=1}^{4}[y_i - (ax_i + b)]^2 = [1 - (0+b)]^2 + [2 - (a+b)]^2 + [4 - (2a+b)]^2 + [8 - (3a+b)]^2.$$

For the minimum to occur at (a, b), it is necessary for

$$\frac{\partial}{\partial a}[E(a,b)] = 0 \quad \text{and} \quad \frac{\partial}{\partial b}[E(a,b)] = 0.$$

Consequently, we obtain a set of equations known as the *Normal Equations.*

$$\begin{cases} a\sum_{i=1}^{4} x_i^2 + b\sum_{i=1}^{4} x_i = \sum_{i=1}^{4} x_i y_i \\ a\sum_{i=1}^{4} x_i + 4b = \sum_{i=1}^{4} y_i \quad (n=4) \end{cases}$$

which simplifies to

$$\begin{cases} 14a + 6b = 34 \\ 6a + 4b = 15 \end{cases}$$

The solution to this system of equations is $a = 2.3$ and $b = 0.3$, so the best linear equation in the least-square sense is

$$y = 2.3x + 0.3\,.$$

Let $E(x) = 2^x - p(x) = 2^x - 2.3x + 0.3$, then $E'(x) = 2^x \ln 2 - 2.3$ and $E''(x) = 2^x(\ln 2)^2$. Thus $E(x)$ is concave-up and its minimum is achieved at $x = \log_2(2.3/\ln 2) = 1.730$. Hence,

$$\max_{-1\le x\le 1}\{|2^x - p(x)|\} = \max_{-1\le x\le 1}\{|2^x - 2.3x - 0.3|\} = \max\{|E(0)|, |E(1.730)|, |E(3)|\} = |E(1.730)| \approx 0.962$$

The Taylor polynomials of degree one that approximate $f(x)$ about x_i's are

$$q_i(x) = 2^{x_i} + (x - x_i)(2^{x_i}\ln 2) = 2^{x_i} + (x - x_i)2^{x_i}0.693.$$

It is not difficult to verify that in the interval $[0,3]$, $p(x) = 2.3x+0.3$ is a better approximation of $f(x) = 2^x$ than

$$q_1(x) = 1 + 0.693x, \quad q_2(x) = 2 + (x-1)0.693, \quad q_3(x) = 4 + (x-2)0.693, \quad \text{and} \quad q_4(x) = 8 + (x-3)0.693.$$

This means that

$$0.962 \approx \max_{0\le x\le 3}\{|f(x) - p(x)|\} < \max_{0\le x\le 3}\{|f(x) - q_i(x)|\} \quad (i = 1,2,3,4)$$

Note. We may use the discrete least-square method to approximate $f(x)$ with functions other than polynomials such as $g(x) = ae^{bx} + c$.

♣ **Least-Square Approximation**. Suppose $f(x) \in C[a,b]$ and that a polynomial of degree at most n, $P_n(x)$ is required that will minimize the error

$$E_n = \int_a^b [f(x) - p_n(x)]^2 dx.$$

To determine a least-square approximating polynomial, that is, a polynomial to minimize E_n, let

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

and define

$$E(a_0, a_1, \ldots, a_n) = \int_a^b \left[f(x) - \sum_{k=0}^{n} a_k x^k \right]^2 dx.$$

The problem is to find real coefficients $a_0, a_1, \ldots, a_n$ to minimize E is that

$$\frac{\partial E}{\partial a_j} = 0 \quad \text{for each } j = 0, 1, \ldots, n.$$

Since

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

$$\frac{\partial E}{\partial a_j} = -2\int_a^b x^j f(x) dx + 2\sum_{k=0}^{n} a_k \int_a^b x^{j+k} dx.$$

Hence, to find $p_n(x)$, the $(n+1)$ linear equations

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

must be solved for the $(n+1)$ unknowns a_j, $j = 0, 1, \ldots, n$. These equations are called the *Normal Equations*. It can be shown that the normal equations always have a unique solution for the continuous function $f(x)$ defined on the interval $[a, b]$.

Example. Find the least-square approximating polynomial of degree one for the function $f(x) = e^x$ on the interval $[-1, 1]$.

Let $p(x) = a_0 + a_1 x$. We need to minimize

$$E(a_0, a_1) = \int_{-1}^{1} (e^x - a_0 - a_1 x)^2 dx.$$

To find a minimum, we set

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

The normal equations for $p(x) = ax + b$ are given by:

$$a_0 \int_{-1}^{1} 1dx + a_1 \int_{-1}^{1} xdx = \int_{-1}^{1} e^x dx$$

$$a_0 \int_{-1}^{1} xdx + a_1 \int_{-1}^{1} x^2 dx = \int_{-1}^{1} xe^x dx$$

Performing the integration yields $a_0 = 1.175$ and $a_1 = 1.104$, so the best linear equation is

$$p(x) = 1.104x + 1.175$$

Let $E(x) = e^x - p(x) = e^x - 1.104x + 1.175$, then $E'(x) = e^x - 1.104$ and $E''(x) = e^x$. Thus $E(x)$ is concave-up and its minimum is achieved at $x = \ln 1.104 = 0.099$. Hence,

$$\max_{-1 \le x \le 1} \{|e^x - p(x)|\} = \max_{-1 \le x \le 1} \{|e^x - 1.104x - 1.175|\} = e - p(1) \approx 0.439$$

The Taylor polynomial of degree one that approximates e^x about $x = 0$ is $q(x) = 1 + x$ with

$$\max_{-1 \le x \le 1} \{|e^x - q(x)|\} = \max_{-1 \le x \le 1} \{|e^x - 1 - x|\} = e - q(1) \approx 0.718$$

♣ **Orthogonal Polynomials**. An integrable function $w(x)$ is called a weight function on $[a,b]$ if $w(x) \geq 0$ for $x \in [a,b]$, but $w(x) \neq 0$ on any subinterval of $[a,b]$.

The purpose of a weight function is to assign 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 $(-1,10$ and more emphasis when $|x|$ is near one. The following are the weight functions of most interest in the developments of the approximation theory:

$$\begin{aligned} w(x) =& 1 \quad a \leq x \leq b \\ w(x) =& \frac{1}{\sqrt{1-x^2}} \quad -1 \leq x \leq 1 \\ w(x) =& e^{-x} \quad 0 \leq x \leq \infty \\ w(x) =& e^{-x^2} \quad -\infty \leq x \leq \infty \,. \end{aligned}$$

Let $w(x)$ be a weight function and define the *Inner Product* of two continuous functions $f(x)$ and $g(x)$ on the interval $[a,b]$ by

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

We say that $f(x)$ and $g(x)$ are *Orthogonal with respect to the weight function* $w(x)$, if we have $< f(x), g(x) >= 0$.

A set $\{\Phi_0(x), \Phi_1(x), \ldots, \Phi_n(x)\}$ is said to be an *Orthogonal set of functions*, for the interval $[a,b]$ if

$$< \Phi_j(x), \Phi_k(x) >= \int_a^b w(x) \Phi_j(x) \Phi_k(x) dx = \alpha_k \delta_{j,k}$$

where $\alpha_k > 0$ and $\delta_{ij} = 1$ whenever $i = j$ and zero otherwise. If, in addition, $\alpha_k = 1$ for each $k = 0, 1, \ldots, n$, the set is said to be *orthonormal.*

Gram-Schmidt (polynomial) Theorem. *Let $\{P_0(x), P_1(x), \ldots, P_n(x)\}$ be set of polynomials defined on $[a,b]$ with degree $P_k(x) = k$ for all k. Then there exists a sequence of orthonormal polynomials $\{\phi_0(x), \phi_1(x), \ldots, \phi_n(x)\}$ on $[a,b]$ corresponding to the polynomials $P_k(x)$.*

♡ **Legendre Polynomials**. Let $w(x) = 1$ on $[-1,1]$. Define

$$P_n(x) = \frac{(-1)}{2^n n!} \cdot \frac{d^n}{dx^n} \left[(1-x^2)^n \right] \quad n \geq 1$$

with $P_0(x) = 1$. These are orthogonal on $[-1,1]$, degree $P_n(x) = n$, and $P_n(1) = 1$ for all n. Also

$$< P_n(x), P_n(x) >= \frac{2}{2n+1}$$

$$\phi_n(x) = \sqrt{\frac{2n+1}{2}} P_n(x).$$

Finally, we have the following *Triple recursion relation*

$$P_{n+1}(x) = \frac{2n+1}{n+1} x P_n(x) - \frac{n}{n+1} P_{n-1}(x).$$

♡ **Chebyshev Polynomials**. Let $w(x) = 1/\sqrt{1-x^2}$ on $[-1,1]$. Then

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

is an orthogonal family of polynomials with degree $T_n(x) = n$. To see this more clearly, let $\arccos(x) = \theta$, or $x = \cos\theta$, $0 \leq \theta \leq \pi$. Then

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

Thus

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

Also by direct calculation, we have

$$T_0(x) = 1 \quad T_1(x) = x$$

Using the above formulae we obtain

$$T_2(x) = 2x^2 - 1 \quad \text{and} \quad T_3(x) = 4x^3 - 3x$$

The polynomials also satisfy

$$T_n(1) = 1 \qquad < T_n(x), T_n(x) >= \begin{cases} \pi, & n = 0 \\ \pi/2 & n > 1 \end{cases}$$

♡ **Laguerre Polynomials**. Let $w(x) = e^{-x}$ on $[0, \infty]$. Then

$$L_n(x) = \frac{1}{n!e^{-x}} \cdot \frac{d^n}{dx^n}\left[x^n e^{-x}\right] \quad n \geq 0$$

is an orthonormal set. and

$$L_{n+1}(x) = \frac{1}{n+1}(2n+1-x)L_n(x) - \frac{n}{n+1}L_{n-1}(x).$$

♣ **Economization of Taylor Polynomials**. Although Taylor polynomials are often very easy to obtain, they are usually inefficient approximations when compared with the minimax approximation of the same degree. Economization is a technique in which first a Taylor polynomial $P_n(x)$ is chosen to yield a certain desired accuracy, say ϵ:

$$|f(x) - P_n(x)| \leq \epsilon \quad a \leq x \leq b$$

This Taylor polynomial is then changed to a polynomial of lower degree without reducing by much the accuracy of $P_n(x)$. We illustrate the process with the following example.

Example. The function $f(x) = e^x$ can be approximated on the interval $[-1,1]$ by the second degree Taylor polynomial $p_2(x) = 1 + x + \frac{1}{2}x^2$ with an error

$$\max_{-1\leq x\leq 1}\{|f(x) - P_2(x)|\} \approx 0.218$$

Notice that

$$\max_{-1\leq x\leq 1}\{|f(x) - P_2(x)|\} \approx 0.218 < R_2(e^x) = \max_{-1\leq x\leq 1}\left\{\frac{|f^{(3)(x)}|}{6}\right\} = \frac{e}{6} \approx 0.453\,.$$

Economization to a degree one polynomial gives the new polynomial

$$M_{2,1}(x) = p_2(x) - \frac{1}{4}T_2(x) = [1 + x + \frac{x^2}{2}] - [\frac{1}{4}(2x^2 - 1)] = x + \frac{5}{4}$$

$$\max_{-1\leq x\leq 1}\{|p_2(x) - M_{2,1}(x)|\} \approx 0.25 \qquad \max_{-1\leq x\leq 1}\{|f(x) - M_{2,1}(x)|\} \leq).218 + 0.25 \approx 0.468$$

The approximation $M_{2,1}(x)$ is linear and is a considerable improvement on the Taylor approximation $P_1(x)$ for which the maximum error is 0.718.

We will let $M_{r,s}(x)$ denote the polynomial obtained by reducing a Taylor polynomial of degree r down to a polynomial of degree s, one degree at a time using the above economization procedure.

Again let $f(x) = e^x$, and produce $p_3(x) = 1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3$. Then

$$\max_{-1\leq x\leq 1}\{|f(x) - P_3(x)|\} \approx 0.0516$$

We will perform economization on $p_3(x)$ twice to obtain a linear approximation.

$$\begin{aligned} M_{3,2}(x) =& p_3(x) - \frac{1}{24}T_3(x) = 1 + \frac{9}{8}x + \frac{1}{2}x^2 \\ M_{2,1}(x) =& M_{3,2}(x) - \frac{1}{4}T_2(x) = \frac{5}{4} + \frac{9}{8}x. \end{aligned}$$

By direct computation,

$$\max_{-1\leq x\leq 1}\{|f(x) - M_{3,1}(x)|\} \approx 0.34$$

a very good linear approximation error.